OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: July 10, 2004, 08:26:09 ; Search time 4183 Seconds
                                       (without alignments)
                                       10258.093 Million cell updates/sec

Title: US-09-972-032-1
Perfect score: 990
Sequence: 1 ggaattgttctcgaggccaa........aaaatgagcggccgcgcaagtt 990

Scoring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151699 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
        1:   gb_ba:*
        2:   gb_htg:*
        3:   gb_in:*
        4:   gb_om:*
        5:   gb_ov:*
        6:   gb_pat:*
        7:   gb_ph:*
        8:   gb_pl:*
        9:   gb_pr:*
       10:   gb_ro:*
       11:   gb_sts:*
       12:   gb_sy:*
       13:   gb_un:*
       14:   gb_vi:*
       15:   em_ba:*
       16:   em_fun:*
       17:   em_hum:*
       18:   em_in:*
       19:   em_mu:*
       20:   em_om:*
       21:   em_or:*
       22:   em_ov:*
       23:   em_pat:*
       24:   em_ph:*
       25:   em_pl:*
       26:   em_ro:*
       27:   em_sts:*
       28:   em_un:*
       29:   em_vi:*
       30:   em_htg_hum:*
       31:   em_htg_inv:*
       32:   em_htg_other:*
       33:   em_htg_mus:*
       34:   em_htg_pln:*
       35:   em_htg_rod:*
       36:   em_htg_mam:*
       37:   em_htg_vrt:*
       38:   em_sy:*
       39:   em_htgo_hum:*
       40:   em_htgo_mus:*
       41:   em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	797.6	80.6	182230	9	AC135050	AC135050 Homo sapi
	2	792.8	80.1	195476	9	AC135044	AC135044 Homo sapi
c	3	87.4	8.8	1602	6	AX884081	AX884081 Sequence
c	4	87.4	8.8	1602	6	BD160707	BD160707 Primer fo
c	5	87.4	8.8	1602	9	AK024161	AK024161 Homo sapi
c	6	87.4	8.8	32360	9	AC005339	AC005339 Homo sapi
	7	86.6	8.7	176735	2	AC021271	AC021271 Homo sapi
	8	86	8.7	185316	2	AC134386	AC134386 Papio anu
c	9	85.8	8.7	2209	9	BC017672	BC017672 Homo sapi
	10	85.8	8.7	80429	2	AC010603	AC010603 Homo sapi
c	11	85.8	8.7	161267	2	AC008378	AC008378 Homo sapi
	12	85.8	8.7	165762	2	AC124263	AC124263 Homo sapi
	13	85.8	8.7	185378	2	AC118754	AC118754 Homo sapi
	14	85.8	8.7	190466	2	AC022285	AC022285 Homo sapi
	15	85.8	8.7	199463	2	AC129497	AC129497 Homo sapi
c	16	85.6	8.6	101658	9	AL590005	AL590005 Human DNA
	17	85.6	8.6	174578	9	AC013287	AC013287 Homo sapi
	18	85.4	8.6	143065	9	HS20208	HS20208 Human DNA
	19	85.4	8.6	173028	9	AC090371	AC090371 Homo sapi
	20	85.4	8.6	193486	9	AC012123	AC012123 Homo sapi
c	21	85.2	8.6	119841	9	AC005002	AC005002 Homo sapi
c	22	85.2	8.6	123291	9	AC104070	AC104070 Homo sapi
	23	85.2	8.6	174133	9	AC093183	AC093183 Homo sapi
c	24	85.2	8.6	176875	9	AC018638	AC018638 Homo sapi
c	25	85	8.6	60500	9	AC131238	AC131238 Homo sapi
c	26	85	8.6	91692	9	AC084253	AC084253 Homo sapi
c	27	85	8.6	94296	2	AL160261	AL160261 Homo sapi
	28	85	8.6	122556	2	AL391863	AL391863 Human DNA
	29	85	8.6	159933	2	AC009692	AC009692 Homo sapi
c	30	85	8.6	161970	9	AP005660	AP005660 Homo sapi
	31	85	8.6	187268	2	AC094022	AC094022 Homo sapi
	32	85	8.6	200598	2	AC012505	AC012505 Homo sapi
c	33	85	8.6	206533	9	CNS01RH1	CNS01RH1 Human chr
c	34	85	8.6	214866	9	AC069544	AC069544 Homo sapi
c	35	85	8.6	114596	9	HS1063B2	HS1063B2 Human DNA
c	36	84.8	8.6	169197	9	AC129096	AC129096 Papio anu
	37	84.8	8.6	202399	2	AC090966	AC090966 Papio anu
c	38	84.8	8.6	231260	2	AL160172	AL160172 Homo sapi
	39	84.6	8.5	91767	9	AC007748	AC007748 Homo sapi
c	40	84.6	8.5	114950	9	AC097652	AC097652 Homo sapi
c	41	84.6	8.5	134308	9	AC011488	AC011488 Homo sapi
	42	84.6	8.5	159440	9	AC079012	AC079012 Homo sapi
	43	84.6	8.5	171611	2	AC119423	AC119423 Papio anu
	44	84.6	8.5	183301	2	AC092487	AC092487 Homo sapi
	45	84.6	8.5	210608	9	AC006028	AC006028 Homo sapi

ALIGNMENTS

RESULT 1
AC135050
LOCUS   AC135050   182230 bp   DNA   linear   PRI 27-FEB-2003
DEFINITION Homo sapiens chromosome 16 clone RP11-196G11, complete sequence.
ACCESSION   AC135050
VERSION   AC135050.3 GI:28570306
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 182230)
AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
TITLE   Direct Submission

```
JOURNAL    Unpublished
REFERENCE  2  (bases 1 to 182230)
AUTHORS    DOE Joint Genome Institute.
TITLE      Direct Submission
JOURNAL    Submitted (05-OCT-2002) Production Sequencing Facility, DOE Joint
           Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE  3  (bases 1 to 182230)
AUTHORS    DOE Joint Genome Institute.
TITLE      Direct Submission
JOURNAL    Submitted (22-OCT-2002) Production Sequencing Facility, DOE Joint
           Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE  4  (bases 1 to 182230)
AUTHORS    DOE Joint Genome Institute, Stanford Human Genome Center and Los
           Alamos National Laboratory.
TITLE      Direct Submission
JOURNAL    Submitted (27-FEB-2003) DOE Joint Genome Institute, 2800 Mitchell
           Drive, Walnut Creek, CA 94598, USA
COMMENT    On Feb 27, 2003 this sequence version replaced gi:24211094.
           Draft Sequence Produced by DOE Joint Genome Institute
           www.jgi.doe.gov
           Finishing Completed at Stanford Human Genome Center and Los Alamos
           National Laboratory.
           www-shgc.stanford.edu
           Quality: Phrap Quality >=40 99.7% of Sequence;
           Estimated Total Number of Errors is 0.3.
FEATURES            Location/Qualifiers
     source         1..182230
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /chromosome="16"
                    /clone="RP11-196G11"
ORIGIN

Query Match     80.6%;  Score 797.6;  DB 9;  Length 182230;
Best Local Similarity  99.3%;  Pred. No. 1.4e-181;
Matches  822;  Conservative  0;  Mismatches  4;  Indels  2;  Gaps  2;

Qy    126 TTTAAAGGGCCCCTGCGTGTTGCTGCGCGCCGCTCCCAGTGCGGG 185
          ||   ||||||||||||||||||| ||||||||||||||||||||||
Db  98012 TCTTACAGGGCCCCTGTTGTTGCTGCCGCTGCCGCCGCTCCCAGTGCCCAGTCTGCGGGG 98071

Qy    186 CTCAGTCCCGGCTTGCCATGTGGGAGACCGCGTCGCGTAAGCGCTGGATGTGGCTTCG 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98072 CTCAGTCCCGGCTTGCCATGTGGGAGACCGCGTCGCGTAAGCGCTGGATGTGGCTTCG 98131

Qy    246 CTGATGCACATTGGACCGGGCTCTGGACTGGGCTAGGGGAAGGGCAGGAGGGCGAATTG 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98132 CTGATGCACATTGGACCGGGCTCTGGACTGGGCTAGGGGAAGGGCAGGAGGGCGAATTG 98191

Qy    306 GGCCGAGGGCCAGGCCCAGGCCTCGCCGCACCCCGACCCCGCCGCTGGCCCGCAGCGCCT 365
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98192 GGCCGAGGGCCAGGCCCAGGCCTCGCCGCACCCCGACCCCGCCGCTGGCCCGCAGCGCCT 98251

Qy    366 CCCGGTGGCCCTGGAGTGCAGTGCAGTTGCAGGATCGTCCGGAGATCGTCCGGAGCTGT 425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98252 CCCGGTGGCCCTGGAGTGCAGTGCAGTTGCAGGATCGTCCGGAGATCGTCCGGAGCTGT 98311

Qy    426 GCATGGGCGGTGGCTAAGGCCGTGGTTGGTTACGATTGGTTACGATTGGCCAGCGGGAC 485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98312 GCATGGGCGGTGGCTAAGGCCGTGGTTGGTTACGATTGGTTACGATTGGCCAGCGGGAC 98371

Qy    486 TCTCTGAAGAGCATGACATTAGTCTGGAGGGTCCTGGAGAGTGATCCCCGCCCACCA 545
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98372 TCTCTGAAGAGCATGACATTAGTCTGGAGGGTCCTGGAGAGTGATCCCCGCCCCACCA 98431

Qy    546 TCAAATGGCGCTTAGGTCTAGGAAGCGGGTGGGTGGGCCTTAGGGCCTTAGGGCGCAGAC 605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98432 TCAAATGGCGCTTAGGTCTAGGAAGCGGGTGGGTGGGCCTTAGGGCCTTAGGGCGCAGAC 98491

Qy    606 ATACCCCGAAGTGGTTGGATTGTATACCGCAAGGGCTCAAGGGCTCGAACCCCCAAAGACAC 665
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98492 ATACCCCGAAGTGGTTGGATTGTATACCGCAAGGGCTCAAGGGCTCGAACCCCCAAAGACAC 98551

Qy    666 TGGAAGGCTGTGTGGCTGAGGAGGGCCCCGGCA-ATCCAGTGTGTGTGGGCTTTACAGGA 724
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  98552 TGGAAGGCTGTGTGGCTGAGGAGGGCCCCGGCAGATCCCAGTGTGTGGGCTTTACAGGA 98611

Qy    725 AAGAGCTCCACCTCT-TGGAGTGTGCAGATGGCGATCTAGGTGTGTCCACCCATGGGAG 783
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  98612 AAGAGCTCCACTTCTCTGGAGTGTGCAGATGGCGATCTAGGTGTGTCCACCCGATGGGAG 98671

Qy    784 CTGCGGGCCGGCAGATGCGCCCCAGTACAAACGTGATTGGACCTGGGGCCTCTGGAC 843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98672 CTGCGGGCCGGCAGATGCGCCCCAGTACAAACGTGATTGGACCTGGGGCCTCTGGAC 98731

Qy    844 TTCCCTGATTCTCTGCTGCCATCCAGCAAGCAAAGTCCTGTCCCGTTGGCTGCCTTCATCCA 903
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98732 TTCCCTGATTCTCTGCTGCCATCCAGCAAGCAAAGTCCTGTCCCGTTGGCTGCCTTCATCCA 98791

Qy    904 CTCTCTCACTTCTCTGCCTTCAGAGTAAAATTGCAAGATCTGTGGTGC 951
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  98792 CTCTCTCACTTCTCTGCCTTCAGAGTAAAATTGCAAGATCTGTGGTGC 98839


RESULT 2
AC135044
LOCUS       AC135044                195476 bp    DNA     linear   HTG 05-OCT-2002
DEFINITION  Homo sapiens chromosome 16 clone CTD-2551B20, WORKING DRAFT
            SEQUENCE, 22 unordered pieces.
ACCESSION   AC135044
VERSION     AC135044.1  GI:23505535
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 195476)
AUTHORS     DOE Joint Genome Institute.
TITLE       Sequencing of Human Chromosome 16
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 195476)
AUTHORS     DOE Joint Genome Institute.
TITLE       Direct Submission
JOURNAL     Submitted (05-OCT-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -----------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            --------
            Project Information
            Center Project Name: 809609
            Center clone name: CITB-E1_2551B20
            --------
            Summary Statistics
            Consensus quality: 171229 bases at least Q40
            Consensus quality: 182638 bases at least Q30
            Consensus quality: 188095 bases at least Q20
            Estimated insert size: 170000; agarose-fp estimation
            Estimated insert size: 193376; sum-of-contigs estimation
            Quality coverage:  7.63 in Q20 bases; agarose-fp estimation
            Quality coverage:  6.7 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 22 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *      1    1202: contig of 1202 bp in length
            *   1203    1302: gap of unknown length
            *   1303    2468: contig of 1166 bp in length
            *   2469    2568: gap of unknown length
            *   2569    4077: contig of 1509 bp in length
```

```
Db  106654 GCATGGGGCGTGGCTAAGGCGTGGTTGGTTACGATTGGCCAGCGGGACTTAAGTGTTG 106713
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     486 TCTCTGAAGAGCATGACATTAGTCTGGAGGGTCCTGGAGAGTGATCCCGCCACCA     545
           |||||||||||||||||||||||||||| ||||||||| |||||||||||||||||
Db  106714 TCTCTGAAGAGCATGACATTAGTCTGGAGAGGTCCTGGAAGAGTGATCCCGCCACCA 106773
Qy     546 TCAAATGGCGCTTAGGTCTAGGAAGCGGTGTGGGTGGGGGCCTTAGGGCGGCAGAC    605
           ||||||||||||||||||||||||||||  ||||||||||||||||||||||||  |
Db  106774 TCAAATGGCGCTTAGGTCTAGGAAGCGGGGGTGTGGGTGGGGGCCTTAGGGCGGCGAGAC 106833
Qy     606 ATACCCGAAGTGGTTGGATTGTATACCGGAAGGGCTGGATCGAACCCCCAAAGACAC   665
           | |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  106834 ACACCCGAAGTGGTTGGATTGTATACCGGAAGGGCTGGATCGAACCCCCAAAGACAC 106893
Qy     666 TGGAAGGCTGTGTGGCTGGAGGAGGGCCGGCA-ATCCAGTGTGTCGTTTACAGGA     724
           ||||||||||||||||||||||||||||||| | ||||||||||||| |||||||
Db  106894 TGGAAGGCTGTGTGGCTGGAGGAGGGCCGGCGAGATCCAGTTCAGTGTGTCGTCCTGGGAG 106953
Qy     725 AAGAGCTCCACCTTCT-TGGAGTGTGCAGATGCGATCTAGGTGTGTCCACCGGATGGGAG   783
           ||||||||||||||| | |||||||||||||||||||||||||||||||||||||||||
Db  106954 AAGAGCTCCACCTTCTTCTGGAGTGTGCAGATGCGATCTAGGTGTGTCCACCGGATGGGAG 107013
Qy     784 CTGCGGGGCCGGGCAGATGCGTCGCCCAGTACAAAGCTGATTTGGACCTGGGCCTCTGGAC  843
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  107014 CTGCGGGGCCGGGCAGATGCGTCGCCCAGTACAAAGCTGATTTGGACCTGGGCCTCTGGAC 107073
Qy     844 TTCCCTGATTCTCTGCTTGCATCTCCAGCAAAGTCCTGTCCGTTGGCTTGGCTGCCTTCATCCA  903
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  107074 TTCCCTGATTCTCTGCTTGCATCTCCAGCAAAGTCCTGTCCGTTGGCTTGGCTGCCTTCATCCA 107133
Qy     904 CTCTCTCACTTCTCTGCCTTCAGAGTAAAATTGCAAGATCTGTGGTGC           951
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db  107134 CTCTCTCACTTCTCTGCCTTCAGAGTAAAATTGCAAGATCTGTGGTGC 107181
```

```
RESULT 3
AX884081/c
LOCUS       AX884081                1602 bp    DNA      linear   PAT 17-DEC-2003
DEFINITION  Sequence 18986 from Patent EP1074617.
ACCESSION   AX884081
VERSION     AX884081.1  GI:40038982
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 18986 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1..1602
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
Query Match             8.8%;  Score 87.4;  DB 6;  Length 1602;
Best Local Similarity  85.8%;  Pred. No. 1.9e-10;
Matches    97;  Conservative   0;  Mismatches  16;  Indels  0;  Gaps  0;

Qy      34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCCAGCTTCCCAAAGTGCTGGATTACAGGCA     93
           ||||||||||| ||||||||||| ||||| |||||||| |||| |||||| ||||||||
Db     935 CGATCTCCTGACCTCGTGATCCGCCCGCCTCCAGCTTCCCAAAGTGCTGGATTACAGGCG    876
Qy      94 TGAGCCACCGCGCCCGGCCCCCGACCTAGTTTAAAGGGCCCCTGCTGTG            146
           ||||||||||||||| |||||||| ||| ||||| |||||||||||
Db     875 TGAGCCACCGCGCCCGGCCCCCGCCCCAGCCCCTTTTTTTAAGGACCCCCAGTCCTTG    823
```

FEATURES
```
*   4078  4177: gap of unknown length
*   4178  5772: contig of 1595 bp in length
*   5772  5872: gap of unknown length
*   5873  7722: contig of 1850 bp in length
*   7722  7822: gap of unknown length
*   7823  8958: contig of 1136 bp in length
*   8959  9058: gap of unknown length
*   9059 11051: contig of 1993 bp in length
*  11052 11151: gap of unknown length
*  11152 13641: contig of 2490 bp in length
*  13642 13741: gap of unknown length
*  13742 17738: contig of 3997 bp in length
*  17739 17838: gap of unknown length
*  17839 22030: contig of 4192 bp in length
*  22031 22130: gap of unknown length
*  22131 27002: contig of 4872 bp in length
*  27003 27102: gap of unknown length
*  27103 32819: contig of 5717 bp in length
*  32820 32919: gap of unknown length
*  32920 39003: contig of 6084 bp in length
*  39004 39103: gap of unknown length
*  39103 48370: contig of 9267 bp in length
*  48371 48470: gap of unknown length
*  48471 59721: contig of 11251 bp in length
*  59722 59821: gap of unknown length
*  59822 72304: contig of 12483 bp in length
*  72305 72404: gap of unknown length
*  72405 85349: contig of 12945 bp in length
*  85350 85449: gap of unknown length
*  85450 101872: contig of 16423 bp in length
* 101873 101972: gap of unknown length
* 101973 117076: contig of 15104 bp in length
* 117077 117176: gap of unknown length
* 117177 138973: contig of 21797 bp in length
* 138974 139073: gap of unknown length
* 139074 150336: contig of 11263 bp in length
* 150337 150436: gap of unknown length
* 150437 195476: contig of 45040 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..195476
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="CTD-2551B20"
                     /clone_lib="CalTech human BAC library D"
ORIGIN
```

```
Query Match            80.1%;  Score 792.8;  DB 2;  Length 195476;
Best Local Similarity  98.9%;  Pred. No. 2e-180;
Matches   819;  Conservative   0;  Mismatches   7;  Indels   2;  Gaps   2;

Qy     126 TTTAAAGGGCCCCTGCTGTTGCTGCCGCTCCCAGCTGCCCAGCTGCCCAGTCTGCGCGGG    185
           |||| |||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db  106354 TCTTACAGGGCCCCTGCTGTTGCTGCCGCTCCCAGCTGCCCAGCTGCCCAGTCTGCGCGGG 106413
Qy     186 CTCAGTCCGCGGTTGCCATGTGGGGAGACCCGGTCGGGTAAGCGCTGGATGTGGCTTCG    245
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  106414 CCCAGTCCGCGGTTGCCATGTGGGGAGACCCGGTCGGGTAAGCGCTGGATGTGGCTTCG 106473
Qy     246 CTGATGCACATTGGACCGGGCTTCTGGACTGGGCTAGGGAGGGCAGGAGGGCCGGAATTG    305
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  106474 CTGATGCACATTGGACCGGGCTTCTGGACTGGGCTAGGGAGGGCAGGAGGGCCGGAATTG 106533
Qy     306 GGCCCGAGGGCCAGGCCTTGCCCGGACACCCGGTCGGGCTCCCGGTGGCCCCGCAGCGGGCT    365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  106534 GGCCCGAGGGCCAGGCCTTGCCCGGACACCCGGTCGGGCTCCCGGTGGCCCCGCAGCGGGCT 106593
Qy     366 CCCGGTGGCCCTGGAGTGCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGGCGAGCTGT    425
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  106594 CCCGGTGGCCCTGGAGTGCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGGCGAGCTGT 106653
Qy     426 GCATGGGGCGTGGCTAAGGCGTGGTTGGTTACGATTGGCCAGCGGGACTTAAGTGTTG    485
```

```
RESULT 4
BD160707/c
LOCUS       BD160707      1602 bp   DNA    linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD160707
VERSION     BD160707.1  GI:27866465
KEYWORDS    JP 2002191363-A/15550
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1602)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 15550 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/15550
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC   C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N5/ PC
            10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
            Primer for synthesizing full-length cDNA and use thereof FH   Key
            FT   source          1..1602
            FT                   /organism='Homo sapiens (human)' .
FEATURES    source          1..1602
                            /organism="Homo sapiens"
                            /mol_type="genomic DNA"
                            /db_xref="taxon:9606"
ORIGIN

Query Match             8.8%;    Score 87.4;  DB 6;  Length 1602;
Best Local Similarity  85.8%;    Pred. No. 1.9e-10;
Matches   97;  Conservative    0;  Mismatches  16;  Indels    0;  Gaps    0;

Qy     34 CGATCTCCTGACCTCGTGATCCGCCGCCCGGCCTCCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    935 CGATCTCCTGACCTCGTGATCCGCCGCCCGGCCTCCAGCCTCTCAAAGTGCTGGGATTACAGGCG 876

Qy     94 TGAGCCACCGCGCCCGGCCCGACACCTAGTTTAAAGGGGCCCCTGCTGTTG 146
          |||||||||||||||||||||| |||| ||||| |||| || |||
Db    875 TGAGCCACCGCGCCCGGCCCTGCCCCAGGCCCCTTTTTTTAAGGACCCCAGTCCTTG 823


RESULT 5
AK024161/c
LOCUS       AK024161      1602 bp   mRNA   linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens cDNA FLJ14099 fis, clone MAMMA1000802.
ACCESSION   AK024161
VERSION     AK024161.1  GI:10436472
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1602)
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
            Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
            Arita,M., Nabekura,T., Ishii,S., Kawai,F., Saito,K., Yamamoto,J.,
            Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y. and Oshima,A.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1602)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp; Tel:81-438-52-3975; Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:
            Helix Research Institute (supported by Japan Key Technology Center
            etc.) and Department of Virology, Institute of Medical Science,
            University of Tokyo.
FEATURES    Location/Qualifiers
     source 1..1602
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="MAMMA1000802"
            /tissue_type="Mammary gland"
            /clone_lib="MAMMA"
            /note="cloning vector: pME18SFL3"
ORIGIN

Query Match             8.8%;    Score 87.4;  DB 9;  Length 1602;
Best Local Similarity  85.8%;    Pred. No. 1.9e-10;
Matches   97;  Conservative    0;  Mismatches  16;  Indels    0;  Gaps    0;

Qy     34 CGATCTCCTGACCTCGTGATCCGCCGCCCGGCCTCCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    935 CGATCTCCTGACCTCGTGATCCGCCGCCCGGCCTCCAGCCTCTCAAAGTGCTGGGATTACAGGCG 876

Qy     94 TGAGCCACCGCGCCCGGCCCGACACCTAGTTTAAAGGGGCCCCTGCTGTTG 146
          |||||||||||||||||||||| |||| ||||| |||| || |||
Db    875 TGAGCCACCGCGCCCGGCCCTGCCCCAGGCCCCTTTTTTTAAGGACCCCAGTCCTTG 823


RESULT 6
AC005339/c
LOCUS       AC005339      32360 bp   DNA    linear   PRI 30-JUL-1998
DEFINITION  Homo sapiens chromosome 19, cosmid R33729, complete sequence.
ACCESSION   AC005339
VERSION     AC005339.1  GI:3355454
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 32360)
  AUTHORS   Lamerdin,J.E., McCready,P.M., Skowronski,E., Adamson,A.W.,
            Burkhart-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stilwagen,S.,
            Phan,H., Velasco,N. Do,L., Regala,W., Terry,A., Garnes,J.,
            Danganan,L., Poundstone,P., Christensen,M., Georgescu,A., Avila,J.,
            Liu,S., Attix,C., Andreise,T., Trankheim,M., Amico-Keller,G.,
            Coefield,J., Duarte,S., Lucas,S., Bruce,R., Thomas,P., Quan,G.,
            Kronmiller,B., Arellano,A., Montgomery,M., Ow,D., Nolan,M.,
            Trong,S., Kobayashi,A., Olsen,A.S. and Carrano,A.V.
  TITLE     Sequence analysis of a 1 Mb region in 19p13.3
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 32360)
  AUTHORS   Lamerdin,J.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-1998) Joint Genome Institute, Lawrence Livermore
            National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
COMMENT     Map and sequence oriented from p telomere to centromere. Cosmid
            R33729 overlaps BAC 48708 to the left from bases 1 to 8,574 and
            overlaps cosmid R26894 to the right from bases 26,845 to 32,360.
            Additional chr 19 map and sequence information are available at:
            http://www-bio.llnl.gov/genome/genome.html.
FEATURES    Location/Qualifiers
     source 1..32360
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
```

```
                /chromosome="19"
                /map="19p13.3 between D19S883 and MLLT1"
                /clone="R33729"
                /cell_line="5HL2-B"
                /clone_lib="LL19NC03 R chromosome 19-specific cosmid
                library"
                /note="Cosmid library constructed at LLNL from flow-sorted
                chromosomes from human-hamster hybrid 5HL2-B, which
                carries chromosome 19 as its only human chromosome."
repeat_region   1..249
                /rpt_family="AluSx"
repeat_region   complement(354..693)
                /rpt_family="LINE2"
repeat_region   710..787
                /rpt_family="MIR"
repeat_region   complement(1403..1521)
                /rpt_family="LINE2"
repeat_region   1562..1672
                /rpt_family="AluJb"
repeat_region   1689..1780
                /rpt_family="(GA)n"
repeat_region   complement(2197..2502)
                /rpt_family="AluSx"
repeat_region   3589..3609
                /rpt_family="GC_rich"
repeat_region   4266..4572
                /rpt_family="AluJo"
repeat_region   4692..4784
                /rpt_family="MIR"
repeat_region   4793..4874
                /rpt_family="MIR"
repeat_region   complement(4888..5232)
                /rpt_family="AluJo"
repeat_region   complement(5239..5545)
                /rpt_family="AluSx"
misc_feature    5741..5848
                /note="predicted exon, program: grail2exons_human_1.3,
                frame: 1, quality: good, score: 61.000"
repeat_region   complement(6054..6181)
                /rpt_family="FLAM_C"
repeat_region   complement(6188..6262)
                /rpt_family="MER5A"
repeat_region   6677..6826
                /rpt_family="MIR"
repeat_region   6826
                /rpt_family="GC_rich"
repeat_region   complement(6973..7032)
                /rpt_family="(CAGA)n"
repeat_region   7764..7895
                /rpt_family="MIR"
repeat_region   7919..8022
                /rpt_family="LINE2"
repeat_region   8827..8868
                /rpt_family="MIR"
repeat_region   9104..9184
                /rpt_family="MIR"
repeat_region   9219..9441
                /rpt_family="AluJb"
repeat_region   9444..9743
                /rpt_family="AluSx"
repeat_region   complement(9773..10218)
                /rpt_family="LINE2"
repeat_region   10246..10538
                /rpt_family="AluSx"
repeat_region   10911..11179
                /rpt_family="AluSx"
repeat_region   11184..11484
                /rpt_family="AluSp"
repeat_region   11503..11807
                /rpt_family="AluJo"
repeat_region   complement(11856..12145)
                /rpt_family="AluSx"
repeat_region   12511..12840
                /rpt_family="AluSx"
repeat_region   12868..12968
                /rpt_family="AluJ/FLAM"
repeat_region   13135..13228
                /rpt_family="LINE2"
repeat_region   complement(13408..13671)
                /rpt_family="AluSq"
repeat_region   complement(13674..13880)
                /rpt_family="L1MB7"
repeat_region   complement(13881..14172)
                /rpt_family="AluY"
repeat_region   complement(14175..14555)
                /rpt_family="L1MB7"
repeat_region   complement(14560..14854)
                /rpt_family="AluSx"
repeat_region   complement(14865..15172)
                /rpt_family="AluJo"
repeat_region   15228..15275
                /rpt_family="LINE2"
repeat_region   complement(16172..16474)
                /rpt_family="AluSq"
repeat_region   complement(16664..16716)
                /rpt_family="MIR"
repeat_region   16853..16917
                /rpt_family="MIR"
repeat_region   18003..18246
                /rpt_family="AluY"
repeat_region   complement(18247..18276)
                /rpt_family="AT_rich"
repeat_region   18551..18586
                /rpt_family="AT_rich"
repeat_region   complement(18596..18711)
                /rpt_family="FLAM_C"
repeat_region   complement(18725..18991)
                /rpt_family="AluSq"
misc_feature    19115..19675
                /note="predicted exon, program: grail2exons_human_1.3,
                frame: 1, quality: excellent, score: 92.000"
repeat_region   20087..20262
                /rpt_family="MLT1C"
repeat_region   20269..20558
                /rpt_family="AluY"
repeat_region   20579..20711
                /rpt_family="AluJ"
repeat_region   20720..21034
                /rpt_family="AluSc"
repeat_region   21049..21201
                /rpt_family="AluJ"
repeat_region   21225..21398
                /rpt_family="MLT1C"
misc_feature    complement(21329..21549)
                /note="BLASTN similarity to Z42385 (1..220); match: 0.97;
                score: 9.7e-83; database searched: est; H. sapiens partial
                cDNA sequence"
misc_feature    21389..21554
                /note="predicted exon, program: grail2exons_human_1.3,
                frame: 1, quality: excellent, score: 77.000"
repeat_region   complement(21587..21852)
                /rpt_family="AluJb"
repeat_region   22809..22929
                /rpt_family="AluJ"
repeat_region   22943..23243
                /rpt_family="AluSx"
repeat_region   23246..23404
                /rpt_family="AluJ"
repeat_region   complement(23833..24137)
                /rpt_family="AluJo"
repeat_region   24189..24484
                /rpt_family="AluJb"
misc_feature    complement(24790..25329)
                /note="DDS similarity to overlapping ESTs:
                (25329..25201) AA477269 zu43c12.s1 Scores ovary tumor
                NbHOT Homo sapiens cDNA clone 740758 3'; (401..273); 100%
                identity.~(25329..25201) A477268 zu43c12.r1 Scores ovary
                tumor NbHOT Homo sapiens cDNA clone 740758 5'; (221..349);
```

```
CDS               100% identity.-(25061. .24842) A477268 zu43c12.r1 Soares
                  ovary tumor NbHOT Homo sapiens cDNA clone 740758 5';
                  (350. .5701; 98% identity.-(24792. .25272) N32339 yw82g08.s1
                  Homo sapiens cDNA clone 258782 3'. Score: 925 identity:
                  476/480 (99%) .-(25329. .25028) AA579149 nf28a04.s1
                  NCI_CGAP_Pr1 Homo sapiens cDNA clone IMAGE:915054;
                  (134. .434); 99% identity.-(24790. .25061) AA477269
                  zu43c12.s1 Soares ovary tumor NbHOT Homo sapiens cDNA
                  clone 740758 3'; (272. .1); 100% identity.-Additional EST
                  matches:
                  AA581955, AA467935, AI038745, AI041764, T24716"
                  complement(join(25250. .25329,27176. .27248,27914. .27995,
                  32121. .>32185))
                  /note="Hypothetical partial human protein"
                  /codon_start=1
                  /evidence=not_experimental
                  /product="R33729_1"
                  /protein_id="AAC27824.1"
                  /db_xref="GI:3355455"

Query Match       8.8%;      Score 87.4;  DB 9;  Length 32360;
Best Local Similarity    85.8%;  Pred. No.2e-10;
Matches   97;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         34  CGATCTCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAGTGCTGGGATTACAGGCA  93
               |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      20347  CGATCTCTGACCTCGTGATCCGCCCGCCTCAGCCTCTCAAAGTGCTGGGATTACAGGCG  20288

Qy         94  TGAGCCACCGCCGCCCGGCCCCCGGACCACCTAGTTTTAAAGGGGCCCCTGCTGTTG  146
               ||||||||||||||||||||||||||||||||  |||| |||||||||||||||||||
Db      20287  TGAGCCACCGCCGCCCGGCCCCCTGCCCCAGGCCCCCTTTTTTTAAGGACCCCAGTCCTTG  20235


RESULT 7
AC021271
LOCUS       AC021271              176735 bp    DNA           linear    HTG 10-SEP-2000
DEFINITION  Homo sapiens clone RP11-15H15, WORKING DRAFT SEQUENCE, 31 unordered
            pieces.
ACCESSION   AC021271
VERSION     AC021271.5  GI:110047755
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 176735)
AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens chromosome, clone RP11-15H15
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 176735)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Poilara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 10, 2000 this sequence version replaced gi:7387323.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
```

```
http://ftp.genome.washington.edu/RM/RepeatMasker.html
------ Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
------ Project Information
Center project name: L3461
Center clone name: 15_H15
------ Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 149852 bases at least Q40
Consensus quality: 164690 bases at least Q30
Consensus quality: 170387 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 17735; sum-of-contigs
Quality coverage: 3.3 in Q20 bases; agarose-fp
Quality coverage: 3.4 in Q20 bases; sum-of-contigs
------
* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
------
*              1:     contig of 1130  bp in length
*           1131:     gap of 100 bp
*           1231:     contig of 1314  bp in length
*           2545:     gap of 100 bp
*           2645:     contig of 1106  bp in length
*           3751:     gap of 100 bp
*           3851:     contig of 1111  bp in length
*           4962:     gap of 100 bp
*           5062:     contig of 1381  bp in length
*           6443:     gap of 100 bp
*           6543:     contig of 1342  bp in length
*           7885:     gap of 100 bp
*           7985:     contig of 1336  bp in length
*           9321:     gap of 100 bp
*           9421:     contig of 1401  bp in length
*          10822:     gap of 100 bp
*          10922:     contig of 1127  bp in length
*          12049:     gap of 100 bp
*          12149:     contig of 1031  bp in length
*          13180:     gap of 100 bp
*          13280:     contig of 2039  bp in length
*          15319:     gap of 100 bp
*          15419:     contig of 3695  bp in length
*          19114:     gap of 100 bp
*          19214:     contig of 2971  bp in length
*          22185:     gap of 100 bp
*          22285:     contig of 3757  bp in length
*          26042:     gap of 100 bp
*          26142:     contig of 3706  bp in length
*          29848:     gap of 100 bp
*          29948:     contig of 4839  bp in length
*          34787:     gap of 100 bp
*          34887:     contig of 3904  bp in length
*          38791:     gap of 100 bp
*          38891:     contig of 4571  bp in length
*          43462:     gap of 100 bp
*          43562:     contig of 3898  bp in length
*          47460:     gap of 100 bp
*          47560:     contig of 5729  bp in length
*          53289:     gap of 100 bp
*          53389:     contig of 5928  bp in length
*          59317:     gap of 100 bp
*          59417:     contig of 5018  bp in length
*          64435:     gap of 100 bp
```

```
*    64535    70699: contig of 6165 bp in length
*    70700    70799: gap of 100 bp
*    70800    78000: contig of 7201 bp in length
*    78001    78100: gap of 100 bp
*    85101    85200: gap of 100 bp
*    85201    94940: contig of 9740 bp in length
*    94941    95040: gap of 100 bp
*    95041   105428: contig of 10388 bp in length
*   105429   105528: gap of 100 bp
*   105529   117874: contig of 12346 bp in length
*   117875   117974: gap of 100 bp
*   117975   132905: contig of 14931 bp in length
*   132906   133005: gap of 100 bp
*   133006   150088: contig of 17283 bp in length
*   150089   150388: gap of 100 bp
*   150389   176735: contig of 26347 bp in length.

FEATURES        Location/Qualifiers
source          1..176735
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /clone="RP11-15H15"
                /clone_lib="RPCI-11 Human Male BAC"
misc_feature    1..1130
                /note="assembly_fragment
                clone_end:SP6
                vector_side:left"
misc_feature    1231..2544
                /note="assembly_fragment"
misc_feature    2645..3750
                /note="assembly_fragment"
misc_feature    3851..4961
                /note="assembly_fragment"
misc_feature    5062..6442
                /note="assembly_fragment"
misc_feature    6543..7884
                /note="assembly_fragment"
misc_feature    7985..9320
                /note="assembly_fragment"
misc_feature    9421..10821
                /note="assembly_fragment"
misc_feature    10922..12048
                /note="assembly_fragment"
misc_feature    12149..13179
                /note="assembly_fragment"
misc_feature    13280..15318
                /note="assembly_fragment"
misc_feature    15419..19113
                /note="assembly_fragment"
misc_feature    19214..22184
                /note="assembly_fragment"
misc_feature    22285..26041
                /note="assembly_fragment"
misc_feature    26142..29847
                /note="assembly_fragment"
misc_feature    29948..34786
                /note="assembly_fragment"
misc_feature    34887..38790
                /note="assembly_fragment"
misc_feature    38891..43461
                /note="assembly_fragment"
misc_feature    43562..47459
                /note="assembly_fragment"
misc_feature    47560..53288
                /note="assembly_fragment"
misc_feature    53389..59316
                /note="assembly_fragment"
misc_feature    59417..64434
                /note="assembly_fragment"
misc_feature    64535..70699
                /note="assembly_fragment"
misc_feature    70800..78000
                /note="assembly_fragment"
misc_feature    78101..85100
                /note="assembly_fragment"
misc_feature    85201..94940
                /note="assembly_fragment"
misc_feature    95041..105428
                /note="assembly_fragment"
misc_feature    105529..117874
                /note="assembly_fragment"
misc_feature    117975..132905
                /note="assembly_fragment"
misc_feature    133006..150288
                /note="assembly_fragment"
misc_feature    150389..176735
                /note="assembly_fragment"

ORIGIN

Query Match     8.7%;    Score 86.6;  DB 2;   Length 176735;
Best Local Similarity  91.1%;   Pred. No. 3.2e-10;
Matches 92;  Conservative  0;  Mismatches 9;  Indels 0;  Gaps  0;

Qy    34  CGATCCTCGACCTCGTGATCCGCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA  93
          ||||||||||||||||||||||||||||| ||||| | ||||||||||||||||||||||
Db 42019  CGATCCTCGACCTCGTGATCCGCCGCCTCGGCCTCGGGCCTCCCAAAGTGCTGGGATTACAGGCT 42078

Qy    94  TGAGCCACCGCGCCCGGCGCCCCGACACCTAGTTTTAAAGGG 134
          |||||||||||||||||||||  |||||  ||||||||||||
Db 42079  TGAGCCACCGCGCCCGGCGCCGGCCACTGACAAGCAGTTTTAAATGG 42119

RESULT 8
AC134386
LOCUS       AC134386              185316 bp    DNA     linear   HTG 11-SEP-2003
DEFINITION  Papio anubis clone rp41-10k6, WORKING DRAFT SEQUENCE, 15 ordered
            pieces.
ACCESSION   AC134386
VERSION     AC134386.5  GI:31193964
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Papio anubis (olive baboon)
ORGANISM    Papio anubis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 185316)
AUTHORS     Lau,C.C.Y., Shaikh,T. and Roe,B.A.
TITLE       Papio anubis BAC Clone rp41-10k6
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 185316)
AUTHORS     Lau,C.C.Y. and Roe,B.A.
TITLE       Direct Submission
JOURNAL     Submitted (26-SEP-2003) Department of Chemistry And Biochemistry,
            The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   3  (bases 1 to 185316)
AUTHORS     Lau,C.C.Y., Shaikh,T. and Roe,B.A.
TITLE       Direct Submission
JOURNAL     Submitted (11-SEP-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On May 30, 2003 this sequence version replaced gi:28850063.
            ----------------------- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR

            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 15 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submitter;
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
```

```
*   the accession number will be preserved.
*            1     2467: contig of 2467 bp in length
*         2468     2567: gap of unknown length
*         2568     5707: contig of 3140 bp in length
*         5708     5807: gap of unknown length
*         5808     8278: contig of 2471 bp in length
*         8279     8378: gap of unknown length
*         8379    11812: contig of 3434 bp in length
*        11813    11912: gap of unknown length
*        11913    17899: contig of 5987 bp in length
*        17900    17999: gap of unknown length
*        18000    25937: contig of 7938 bp in length
*        25938    26037: gap of unknown length
*        26038    37583: contig of 11546 bp in length
*        37584    37683: gap of unknown length
*        37684    51903: contig of 14220 bp in length
*        51904    52003: gap of unknown length
*        52004    62301: contig of 10297 bp in length
*        62301    62400: gap of unknown length
*        62401    74061: contig of 11661 bp in length
*        74062    74161: gap of unknown length
*        74162    85046: contig of 10885 bp in length
*        85047    85146: gap of unknown length
*        85147    99131: contig of 13985 bp in length
*        99132    99231: gap of unknown length
*        99232   120762: contig of 21531 bp in length
*       120763   120862: gap of unknown length
*       120863   150205: contig of 29343 bp in length
*       150206   150305: gap of unknown length
*       150306   185316: contig of 35011 bp in length.

FEATURES             Location/Qualifiers
     source          1..185316
                     /organism="Papio anubis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9555"
                     /clone="rp41-10k6"
                     /clone_lib="RPCI - 41 Male (Olive) Baboon BAC Library"
ORIGIN

Query Match          8.7%;   Score 86;   DB 2;   Length 185316;
Best Local Similarity 68.4%;   Pred. No. 4.5e-10;   Indels  0;   Gaps   0;
Matches 119;   Conservative  0;   Mismatches  55;   Indels  0;   Gaps   0;

Qy       34 CGATCTCCTGACCTCGTGATCGCCGCCCTCAGCTTCCAAAGTGCTGGGATTACAGGCA 93
            |||||||||||||||||||||||||||||||||| |||||||| |||||||||||||||
Db   174411 CGATCTCCTGACCTCGTGATCGCCGCCCGTCTCGGCCTCCAAAGTGCTGGGATTACAGGCG 174470

Qy       94 TGAGCCACCGCGCCCGGCCCCGGCCCCCGACACCTAGTTTTAAAGGGCCCCCGTCGTTGCCGC 153
            | ||||||||||||| |||| |||||| |||||||||| |||| ||||||| |||| ||||
Db   174471 TAAGCCACCGCGCCCGGCCCCGCCTCTAATGCTATTTTTAAGCTCATCAGCTATTGTTTAGCGT 174530

Qy      154 TGCCGCCGCTCCAGTGCCAGTCTGGCGGGCTGCCCGGTTGCCATGTG 207
            |||||||||||| || |||| ||||| ||||||||| || ||| |||||||
Db   174531 TAGTGTATTTATGCGTGCGTGCGCCCAAGACAATTTCTTCTTCCAATGTGGCCCAGGG 174584

RESULT 9
BC017672/c
LOCUS       BC017672                2209 bp    mRNA    linear   PRI 06-OCT-2003
DEFINITION  Homo sapiens hypothetical protein MGC17791, mRNA (cDNA clone
            MGC:17791 IMAGE:3885999), complete cds.
ACCESSION   BC017672
VERSION     BC017672.1  GI:17389225
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2209)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
```

```
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   MEDLINE   22388257
    PUBMED   12477992
REFERENCE   2  (bases 1 to 2209)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-DEC-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site:      http://www-shgc.stanford.edu
            Contact:   (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
            R. M.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 22 Row: c Column: 16.
FEATURES             Location/Qualifiers
     source          1..2209
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:17791 IMAGE:3885999"
                     /tissue_type="Lung, carcinoma, large cell
                     undifferentiated "
                     /clone_lib="NIH_MGC_69"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1..2209
                     /gene="MGC17791"
                     /db_xref="LocusID:126282"
     CDS             116..676
                     /codon_start=1
                     /product="hypothetical protein MGC17791"
                     /protein_id="AAH17672.1"
                     /db_xref="GI:17389226"
                     /db_xref="LocusID:126282"
                     /translation="MDTFSTKSLALQAQKLLSKMASKAVVAVLVDDTSSEVLDELYR
                     ATREFTRSRKEAQKMLKNLVKVALIKIGLLLRGDQLGGEELALLRRFRHRARCLAWTAV
                     SFHQVDFTFDRRVLAVGLLECRDLLQAVGPHLTAKSHGRINHVFGHLADCDFLAALY
                     GPAEPYRSHLRRICEGLGRMLDEGSL"
ORIGIN

Query Match          8.7%;   Score 85.8;   DB 9;   Length 2209;
Best Local Similarity 85.0%;   Pred. No. 4.7e-10;   Indels  0;   Gaps   0;
Matches 96;   Conservative  0;   Mismatches 17;   Indels  0;   Gaps   0;

Qy       34 CGATCTCCTGACCTCGTGATCGCCGCCCGGCCCTCAGCTTCCAAAGTGCTGGGATTACAGGCA 93
```

```
Db    1348  ||| |||||||||||||||||||||||||||| ||| |||||||||||||||||||||||||||
            CGACCTCCTGACCTCGTGATCCGCCCGGCGCCCCGACACCTAGTTTCAAAGTGCTGGGATTACAGGCG 1289
Qy      94  TGAGCCACCGCGGCCGGCGCCCCGACACCTAGTTTTAAAGGGCCCTGCTGTTG 146
            ||||||||||||||||||| |||||| |||  ||||      |||| |||
Db    1288  TGAGGCCACCGACCGGCGCCCTGCCCGCCCCAGGCCCCCTTTTTTTAGGACCCCAGTCCTTG 1236
```

RESULT 10
AC010603

```
LOCUS       AC010603        80429 bp  DNA   linear  HTG 18-JUL-2000
DEFINITION  Homo sapiens chromosome 5 clone CTB-137N20, WORKING DRAFT SEQUENCE,
            9 ordered pieces.
ACCESSION   AC010603
VERSION     AC010603.4  GI:9256256
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 80429)
AUTHORS     DOE Joint Genome Institute.
TITLE       Sequencing of Human Chromosome 5
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 80429)
AUTHORS     DOE Joint Genome Institute.
TITLE       Direct Submission
JOURNAL     Submitted (16-SEP-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 18, 2000 this sequence version replaced gi:7710622.
            -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -----
            Project Information
            Center Project Name: 117144
            Center clone name: CIT978SKB_137N20
            -----
            Summary Statistics
            Consensus Quality: 72944 bases at least Q40
            Consensus Quality: 77791 bases at least Q30
            Consensus Quality: 78685 bases at least Q20
            Estimated insert size: 88000; pulse field gel estimation
            Estimated insert size: 80079; sum-of-contigs estimation
            Quality coverage:  5.44 in Q20 bases; pulse field gel estimation
            Quality coverage:  5.98 in Q20 bases; sum-of-contigs estimation.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 9 contigs. Gaps between the contigs
          * are represented as runs of N. The order of the pieces
          * is believed to be correct as given, however the sizes
          * of the gaps between them are based on estimates that have
          * provided by the submittor.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.
          *        1   12855: contig of 12855 bp in length
          *    12856   12955: gap of unknown length
          *    12956   17031: contig of 4076 bp in length
          *    17032   17131: gap of unknown length
          *    17132   36328: contig of 19197 bp in length
          *    36329   36429: gap of unknown length
          *    36430   38360: contig of 1932 bp in length
          *    38361   38460: gap of unknown length
          *    38461   43988: contig of 5528 bp in length
          *    43989   44088: gap of unknown length
          *    44089   51843: contig of 7755 bp in length
          *    51844   51943: gap of unknown length
          *    51944   65558: contig of 13615 bp in length
          *    65559   65658: gap of unknown length
          *    65659   70398: contig of 4740 bp in length
          *    70399   70498: gap of unknown length
          *    70499   80429: contig of 9931 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..80429
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTB-137N20"
                     /clone_lib="CalTech human BAC library B"
ORIGIN

Query Match            8.7%;  Score 85.8;  DB 2;  Length 80429;
Best Local Similarity  74.58%;  Pred. No. 4.9e-10;
Matches 108;  Conservative 0;  Mismatches 37;  Indels 0;  Gaps 0;

Qy      35  GATCTCCTGACCTGTGATCCGCCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCAT 94
            |||||||||||||||||||||||||||||| ||| ||||||||||||||||||||||||
Db   63881  GATCTCCTGACCTCGTGATCCGCCCCGCCTCGGCTCCCAAAGTGCTGGGATTACAGGCAT 63940

Qy      95  GAGCCACCGCGGCCGGCGCCCCACACCTAGTTTTAAAGGGCCCCTGCTGTTGCTGCCGCT 154
            |||||||||||||||||||||       |||||||   ||||| |||  ||
Db   63941  GAGCCACCGCGGCCGGCGCCCCTGAATGTAGTTTTAAATTGTCCCTGCCCACTTTCCTCTTCA 64000

Qy     155  GCCGCGCTCCCAGCTGCCCAGTCT 179
            |||||||||||||||||||
Db   64001  ACTGCAGTTTAGATTGGCCCCTTCT 64025
```

RESULT 11
AC008378/c

```
LOCUS       AC008378        161267 bp  DNA   linear  PRI 20-OCT-2001
DEFINITION  Homo sapiens chromosome 5 clone CTC-209H22, complete sequence.
ACCESSION   AC008378
VERSION     AC008378.6  GI:16303396
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 161267)
AUTHORS     DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 161267)
AUTHORS     DOE Joint Genome Institute.
TITLE       Direct Submission
JOURNAL     Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 161267)
AUTHORS     DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE       Direct Submission
JOURNAL     Submitted (20-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Oct 20, 2001 this sequence version replaced gi:7711255.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.5% of Sequence;
            Estimated Total Number of Errors is 0.5.
FEATURES             Location/Qualifiers
     source          1..161267
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTC-209H22"
ORIGIN

Query Match            8.7%;  Score 85.8;  DB 9;  Length 161267;
Best Local Similarity  74.58%;  Pred. No. 5e-10;
Matches 108;  Conservative 0;  Mismatches 37;  Indels 0;  Gaps 0;

Qy      35  GATCTCCTGACCTGTGATCCGCCCCGCCTCCAGCTTCCCAAAGTGCTGGGATTACAGGCAT 94
```

```
Db     80361  ||||||||   ||||||||||||   |||||||||||||||||||||||||||||||||||||
              GATCTCCTGACCTCGTGATCCCGCGCCCCGACACCTAGTTTAAAGGGCCGCCTTCCCAAAGTGCTGGGATTACAGGCAT  80302
Qy        95  GAGCCACCGCGCCCGGCGCCCCCGACACCTAGTTTAAAGGGCCCCTGCTGTGTTGCTGCCGCT  154
              ||||||||   ||||||||||||  |||||||||||||||||||   |||||||||||||||||||
Db     80301  GAGCCACCGCGCCCGGCGCCCCTGAATGTAGTTTTAAATGTAGTTTTAAATTGTCCCTCTCACTTTCCTCTTTCA  80242

Qy       155  GCGCGCCGCCTCCCAGCTGCCCAGTCT  179
              |||||||||||||||||||||||||||
Db     80241  ACTGCAGTTTAGATTGGCCCCTTCT  80217

RESULT 12
AC124263
LOCUS       AC124263               165762 bp    DNA     linear   HTG 27-MAR-2003
DEFINITION  Homo sapiens chromosome 17 clone RP13-563D20 map 17, 4 unordered
            pieces.
ACCESSION   AC124263
VERSION     AC124263.12  GI:25449560
KEYWORDS    HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 165762)
AUTHORS     Birren,B., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens chromosome 17, clone RP13-563D20
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 165762)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
            Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Nicol,R., Norbu,C., Norman,C.H., O'Connor,J., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
            Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
            Schupbach,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Ye,W.J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (14-JUN-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 165762)
AUTHORS     Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson.R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
```

```
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            *      1     100724: contig of 100724 bp in length
            *     100725  100824: gap of 100 bp
            *     100825  130487: contig of 29663 bp in length
            *     130488  130587: gap of 100 bp
            *     130588  136653: contig of 6066 bp in length
            *     136654  136753: gap of 100 bp
            *     136754  165762: contig of 29009 bp in length.
FEATURES             Location/Qualifiers
     source          1..165762
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP13-563D20"
                     /clone_lib="RPCI-13 Human Female BAC"
ORIGIN
Query Match           8.7%;        Score 85.8;  DB 2;   Length 165762;
Best Local Similarity  80.1%;      Pred. No. 5e-10;
Matches  113;     Conservative    0;  Mismatches  27;   Indels    1;  Gaps   1;

Qy    34  CGATCTCCTGACCTCGTGATCCCGCGCCCGCCTCAGTTCCCAAAGTGCTGGATTACAGGCA  93
          ||||||||||||||||||||||||||||   ||||||   |||||||||||||||||||||
Db 74302  CGATCTCCTGACCTCGTGATCCACCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA  74361

Qy    94  TGAGCCACCGCGCCCGGCCGCCACCCGACCAGTTTTAAAGGGCCCGTGCTGTTGCGCGC  153
          |||||||||||||||||   ||||||||   ||||||||||||||   ||||||||||||||
Db 74362  TGAGCCACCGCGCCCGGCCCGAGCACCACA-TTTTACACGGACTGCCCGTTCCAAAACC  74420

Qy   154  TGCCGCGCCTCCCAGCTGCCC  174
          |||||||||||||| ||||||
Db 74421  TTCTCTGGCTCCCTACTGCCC  74441

RESULT 13
AC118754
LOCUS       AC118754               185378 bp    DNA     linear   PRI 23-JUL-2002
DEFINITION  Homo sapiens, clone RP11-141J13, complete sequence.
ACCESSION   AC118754
VERSION     AC118754.6  GI:21930262
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 185378)
AUTHORS     Birren,B., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens, clone RP11-141J13
```

```
JOURNAL      Unpublished
REFERENCE    2  (bases 1 to 185378)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
             Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
             Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
             Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
             Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
             Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
             Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
             Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,L., Jones,C.,
             Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
             Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
             MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
             McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
             Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
             Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
             Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
             Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
             Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
             Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
             Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
             Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
             Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
             Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE        Direct Submission
JOURNAL      Submitted (20-APR-2002) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE    3  (bases 1 to 185378)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
             Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
             Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
             Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
             Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
             Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
             Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
             Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
             Johnson,L., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
             Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K.,
             Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
             Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
             Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
             Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
             O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
             Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
             Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
             Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
             Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
             Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
             Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
             Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE        Direct Submission
JOURNAL      Submitted (02-JUN-2002) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE    4  (bases 1 to 185378)
AUTHORS      Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
             Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
             Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
             Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
             Gardyna,S., Gord,S., Graham,L., FitzGerald,M., Gage,D., Galagan,J.,
             Horton,L., Hulme,W., Iliev,I., Johnson,L., Jones,C., Kamat,A.,
             Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
             Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
             McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
             Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
             O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
             Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
             Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
             Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
             Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
             Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
             Zembek,L., Zimmer,A. and Zody,M.
TITLE        Direct Submission
```

```
JOURNAL      Submitted (23-JUL-2002) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT      On Jul 23, 2002 this sequence version replaced gi:21313819.
             All repeats were identified using RepeatMasker:
             http://ftp.genome.washington.edu/RM/RepeatMasker.html
             Smit, A.F.A. & Green, P. (1996-1997)
             ----------------- Genome Center
                  Center: Whitehead Institute/ MIT Center for Genome Research
                  Center code: WIBR
                  Web site: http://www-seq.wi.mit.edu
                  Contact: sequence submissions@genome.wi.mit.edu
             ----------------- Project Information
                  Center project name: L26241
                  Center clone name: 141_J_13
             -----------------

FEATURES             Location/Qualifiers
     source          1..185378
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="RP11-141J13"
                     /clone="RPCI-11 Human Male BAC"
     repeat_region   complement(13..297)
                     /rpt_family="AluSp"
     repeat_region   534..577
                     /rpt_family="(CAAAAA)n"
     repeat_region   complement(716..1004)
                     /rpt_family="AluY"
     unsure          730..736
                     /note="<30 qual SNGL region"
     repeat_region   1012..1032
                     /rpt_family="AT_rich"
     repeat_region   complement(1081..1421)
                     /rpt_family="HAL1"
     unsure          1414..1420
                     /note="<30 qual SNGL region"
     repeat_region   1422..1716
                     /rpt_family="AluJo"
     repeat_region   complement(1717..2699)
                     /rpt_family="HAL1"
     unsure          2170..2175
                     /note="<30 qual SNGL region"
     unsure          2207..2211
                     /note="<30 qual SNGL region"
     repeat_region   complement(2700..3006)
                     /rpt_family="AluSg"
     repeat_region   complement(3007..3199)
                     /rpt_family="HAL1"
     repeat_region   complement(3555..3847)
                     /rpt_family="AluY"
     repeat_region   3888..4019
                     /rpt_family="LTR45"
     repeat_region   4087..5497
                     /rpt_family="pTR5"
     repeat_region   complement(5637..5927)
                     /rpt_family="AluSx"
     repeat_region   complement(5934..6242)
                     /rpt_family="AluSp"
     repeat_region   6243..6269
                     /rpt_family="(TCCCC)n"
     repeat_region   6272..6299
                     /rpt_family="(T)n"
     repeat_region   complement(6859..7156)
                     /rpt_family="AluSx"
     repeat_region   7507..7551
                     /rpt_family="(CA)n"
     repeat_region   complement(8312..8550)
                     /rpt_family="AluY"
     repeat_region   complement(8683..8987)
                     /rpt_family="AluSx"
     repeat_region   complement(9065..9469)
                     /rpt_family="MLT1G3"
     repeat_region   9470..9772
```

```
AUTHORS      Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,
             Federspiel,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,
             Mao,J., Komp,C., Kottler,S., Lam,B., Marathe,R., Miranda,M.,
             Morehouse,A.J., Nguyen,M., Oefner,P., Palm,C.J., Ramirez,D.,
             Southwick,A.M., Webb,C., Wilhelmy,J., Yu,S. and Davis,R.W.
  JOURNAL    Unpublished
REFERENCE    2  (bases 1 to 190466)
  AUTHORS    Bruno,D., Conn,L., Dela Rosa,M., Federspiel,N., Foreman,P.,
             Glukhov,S., Hansen,N., Hyman,R., Mao,J., Marathe,R.,
             Morehouse,A.J., Oefner,P., Palm,C.J., Ramirez,D., Wilhelmy,J.,
             Yu,S. and Davis,R.W.
  TITLE      Direct Submission
  JOURNAL    Submitted (28-JAN-2000) DNA Sequencing and Technology Center,
             Stanford University, 855 California Avenue, Palo Alto, CA 94304,
             USA
COMMENT      On Jan 24, 2002 this sequence version replaced gi:16903165.
             -------------- Genome Center
                Center: Stanford DNA Sequencing and Technology Development
             Center
                Center code: SDSTDC
                Web site: http://sequence-www.stanford.edu/group/human/
                Contact: hum-info@sequence.stanford.edu
             -------------- Project Information
                Center project name: 760
                Center clone name: RP11-141J13
             -------------- Summary Statistics
                Sequencing Vector: M13mp18; X02513; 98% of reads
                Sequencing Vector: plasmid; plasmid_accession; 2% of reads
                Chemistry: Dye-primer; 16% of reads
                Chemistry: Dye-terminator Big Dye; 84% of reads
                Assembly program: Phrap; version 0.990319
                Consensus quality: 185754 bases at least Q40
                Consensus quality: 186781 bases at least Q30
                Consensus quality: 187406 bases at least Q20
                Insert size: 184236; agarose-fp
                Insert size: 189766; sum-of-contigs
                Quality coverage: 13.0x in Q20 bases; agarose-fp
                Quality coverage: 12.6x in Q20 bases; sum-of-contigs.
             * NOTE:This is a 'working draft' sequence. It currently
             * consists of 8 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *        1   1958: contig of 1958 bp in length
             *     1959   2058: gap of unknown length
             *     2059   4896: contig of 2838 bp in length
             *     4897   4996: gap of unknown length
             *     4997   8678: contig of 3682 bp in length
             *     8679   8778: gap of unknown length
             *     8779  11477: contig of 2699 bp in length
             *    11478  11577: gap of unknown length
             *    11578  16561: contig of 4984 bp in length
             *    16562  16661: gap of unknown length
             *    16662  39095: contig of 22434 bp in length
             *    39096  39195: gap of unknown length
             *    39196  77173: contig of 37978 bp in length
             *    77174  77273: gap of unknown length
             *    77274 190466: contig of 113193 bp in length.
FEATURES             Location/Qualifiers
     source          1..190466
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /clone="RP11-141J13"
                     /clone_lib="RPCI human BAC library 11"
     misc_feature    1..1958
                     /note="assembly_name:Contig37"
     misc_feature    2059..4896
                     /note="assembly_name:Contig38"
```

```
   unsure                  /rpt_family="AluSc"
                           complement(9688..9707)
                           /note="Single clone coverage"
   repeat_region           9773..9801
                           /rpt_family="(TA)n"
   repeat_region           complement(9835..10132)
                           /rpt_family="AluSx"
   unsure                  complement(9883..9971)
                           /note="single clone coverage"
   unsure                  complement(9915..9931)
                           /note="<30 qual single clone coverage"
   unsure                  complement(9943..9947)
                           /note="<30 qual single clone coverage"
   repeat_region           10162..10771
                           /rpt_family="LIME"
   repeat_region           10781..11096
                           /rpt_family="AluSx"
   repeat_region           complement(11118..11474)
                           /rpt_family="LTR47A"
   repeat_region           11475..11796
                           /rpt_family="AluSx"
   repeat_region           complement(11797..11833)
                           /rpt_family="LTR47A"
   repeat_region           11839..12030
                           /rpt_family="LIME"
   repeat_region           12068..12130
                           /rpt_family="LIME"
   repeat_region           12131..12420
                           /rpt_family="MER89"
   repeat_region           complement(12984..13280)
                           /rpt_family="AluSx"
   repeat_region           complement(13351..13523)
                           /rpt_family="AluJo"
   repeat_region           complement(13524..13831)
                           /rpt_family="AluSx"
   repeat_region           complement(13832..13955)
                           /rpt_family="AluJo"
   repeat_region           14482..14815
                           /rpt_family="AluSx"
   repeat_region           14944..15132
                           /rpt_family="AluJb"
   repeat_region           complement(15166..15605)
```

```
Query Match              8.7%;   Score 85.8;   DB 9;   Length 185378;
Best Local Similarity   80.1%;   Pred. No. 5e-10;
Matches  113;   Conservative   0;   Mismatches  27;   Indels   1;   Gaps    1;

Qy       34  CGATCTCCTGACCTCGTGATCCGCCCGCCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA  93
             ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   142821  CGATCTCCTGACCTCGTGATCCACCCGCCTCAGCCCTCCCAAAGTGCTGGGATTACAGGCG 142880

Qy       94  TGAGCCACCGCGCCCGGCCCCCGACACCTAGTTTTAAAGGGCCCCCTGCTGTTGCTGCCGC 153
             ||||||||||||||||||  ||||||  || |  |||  || | |||||||| |||||||
Db   142881  TGAGCCACCGCGCCCGGCCCCGAGACACACA-TTTTACACGGACTGCCCTGTTCCAAAACC 142939

Qy      154  TGCCGCCGCTCCCAGCTGCCC 174
             |||||||||||||||||||||
Db   142940  TTCTCTGGCTCCCTACTGCCC 142960
```

```
RESULT 14
AC022285
LOCUS       AC022285             190466 bp   DNA     linear   HTG 24-JAN-2002
DEFINITION  Homo sapiens chromosome 17 clone RP11-141J13, WORKING DRAFT
            SEQUENCE, 8 unordered pieces.
ACCESSION   AC022285
VERSION     AC022285.41  GI:18308288
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 190466)
```

```
     misc_feature    4997..8678
                      /note="assembly_name:Contig39"
     misc_feature    8779..11477
                      /note="assembly_name:Contig40
                      clone_end:SP6"
     misc_feature    11578..16561
                      /note="assembly_name:Contig41"
     misc_feature    16662..39095
                      /note="assembly_name:Contig42"
     misc_feature    39196..77173
                      /note="assembly_name:Contig43"
     misc_feature    77274..190466
                      /note="assembly_name:Contig44
                      clone_end:T7"
ORIGIN

Query Match          8.7%;    Score 85.8; DB 2; Length 190466;
Best Local Similarity 80.1%;  Pred. No. 5e-10;
Matches 113;  Conservative  0;  Mismatches  27;  Indels  1;  Gaps     1;

Qy      34  CGATCTCCTGACCTCGTGATCCGCCGCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  147904  CGATCTCCTGACCTCGTGATCCACCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCG 147963

Qy      94  TGAGGCCACCGCGCGCCGCCGCCCCCGACACCTAGTTTAAAGGGCCCCTGCGTGTTGCTGCCGC 153
            |||||||||||||||||  ||| |||||    || | |||| | ||||||||||
Db  147964  TGAGGCCACCGCGCGCCGCCGCCCGAGACACACA-TTTTACACGGACTGCCCTGTTCCAAAACC 148022

Qy     154  TGCCGCCGCTCCCAGCTGCCC 174
            |||||||||||||||||||||
Db  148023  TTCTCTGGCTCCCTACTGCCC 148043


RESULT 15
AC129497
LOCUS       AC129497             199463 bp    DNA      linear    HTG 27-MAR-2003
DEFINITION  Homo sapiens chromosome 17 clone RP11-945P9 map 17, 9 unordered
            pieces.
ACCESSION   AC129497
VERSION     AC129497.3  GI:22758749
KEYWORDS    HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 199463)
AUTHORS     Birren,B., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens chromosome 17, clone RP11-945P9
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 199463)
AUTHORS     Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 199463)
AUTHORS     Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Rogov,P.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Severy,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Talamas,J.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Vassiliev,H.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Young,G., Zainoun,J.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Zody,M.
            Zembek,L., Zimmer,A. and
TITLE       Direct Submission
JOURNAL     Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 8, 2002 this sequence version replaced gi:22475392.
            All repeats were identified using RepeatMasker:
            Smit,A.F.A. & Green,P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----------- Project Information
            Center project name: L27853
            Center clone name: 945_P_9
            -----------
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 9 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *       1     1451:  contig of 1451 bp in length
          *       1452  1551:  gap of 100 bp
          *       1552  5314:  contig of 3763 bp in length
          *       5315  5414:  gap of 100 bp
          *       5415  31286: contig of 25872 bp in length
          *       31287 31386: gap of 100 bp
          *       31387 60741: contig of 29355 bp in length
          *       60742 60841: gap of 100 bp
          *       60842 159763: contig of 98922 bp in length
          *       159764 159863: gap of 100 bp
          *       159864 186613: contig of 26750 bp in length
          *       186614 186713: gap of 100 bp
          *       186714 190447: contig of 3734 bp in length
          *       190448 190547: gap of 100 bp
          *       190548 195762: contig of 5215 bp in length
          *       195763 195862: gap of 100 bp
          *       195863 199463: contig of 3601 bp in length
FEATURES             Location/Qualifiers
     source          1..199463
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-945P9"
                     /clone_lib="RPCI-11 Human Male BAC"
ORIGIN

Query Match          8.7%;    Score 85.8; DB 2; Length 199463;
Best Local Similarity 80.1%;  Pred. No. 5e-10;
Matches 113;  Conservative  0;  Mismatches  27;  Indels  1;  Gaps     1;
```

```
Qy      34  CGATCTCCTGACCTCGTGATCGCCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
                |||||||||||||||||||| |||||||||||| ||||||||||||||||||||||||||
Db  133240  CGATCTCCTGACCTCGTGATCCACCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCG 133299

Qy      94  TGAGCCACCGCGCCCGGCCCCGACACCTAGTTTAAAGGGCCCCTGCTGTTGCTGCCGC 153
                |||||||||||||||| |||||| ||     ||| | |||| ||     |  | 
Db  133300  TGAGCCACCGCGCCCGCCCGGCCCGAGACACACA-TTTTACACGGACTGCCCTGTTCCAAAACC 133358

Qy     154  TGCCGCCGCTCCCAGCTGCCC 174
                |  |||| | |||||| |||||
Db  133359  TTCTCTGGCTCCCTACTGCCC 133379
```

Search completed: July 10, 2004, 11:04:38
Job time : 4193 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2004, 08:24:24 ; Search time 495 Seconds
                                   (without alignments)
                                   8496.396 Million cell updates/sec

Title: US-09-972-032-1
Perfect score: 990
Sequence: 1 ggaattgttctcgaggccaa.........aaaatgagcggcgcaagt 990

Scoring table: IDENTITY_NUC
                  Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_29Jan04:*
         1: geneseqn1980s:*
         2: geneseqn1990s:*
         3: geneseqn2000s:*
         4: geneseqn2001as:*
         5: geneseqn2001bs:*
         6: geneseqn2002s:*
         7: geneseqn2003as:*
         8: geneseqn2003bs:*
         9: geneseqn2003cs:*
        10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	990	100.0	990	6	ABL60606	Abl60606 Human ERC
2	148.6	15.0	597	5	AAS69040	Aas69040 DNA encod
3	87.4	8.8	1602	9	AAH18715	Aah18715 Human cDN
4	84.2	8.5	67782	9	ADC85452	Adc85452 Human Git
5	84.2	8.5	67783	8	ADA02972	Ada02972 Human GIT
6	84.2	8.5	67783	9	ADB72710	Adb72710 Human GIT
7	83.6	8.4	274	3	AAC19185	Aac19185 Human sec
8	83.2	8.4	2105	4	AAH99495	Aah99495 Human pro
9	83.2	8.4	2214	5	AAK66423	Aak66423 Human imm
10	83.2	8.4	6742	4	AAK66424	Aak66424 Human imm
11	83.2	8.4	10772	4	AAK66425	Aak66425 Human imm
12	83	8.4	13695	9	ADC86452	Adc86452 Human GPC
13	82.6	8.3	2251	4	AAH17556	Aah17556 Human cDN
14	82.2	8.3	444	4	AAK85393	Aak85393 Human imm
15	82.2	8.3	2390	9	ADB63236	Adb63236 Human cDN
16	82.2	8.3	12477	4	AAS35051	Aas35051 DNA #1 en
17	82.2	8.3	12477	4	AAK85389	Aak85389 Human imm
18	82.2	8.3	12477	4	AAS31480	Aas31480 Human DNA
19	82.2	8.3	12477	4	AAI63967	Aai63967 Human pol
20	82.2	8.3	12477	4	ABA06765	Aba06765 Human gen
21	82.2	8.3	12477	6	ABQ66804	Abq66804 Human pol
22	82.2	8.3	12477	6	ABV84102	Abv84102 Human pol
23	82.2	8.3	12477	9	ADC11091	Adc11091 Human DNA

24	82.2	8.3	12477	9	ADC46493	Adc46493 Human neo
25	82	8.3	2904	9	ADB63202	Adb63202 Human cDN
26	82	8.3	16491	4	AAL02817	Aal02817 Human rep
27	81.8	8.3	4686	5	ABA18463	Aba18463 Human ner
28	81.6	8.2	387	6	ABL87779	Abl87779 Human ova
29	81.4	8.2	1502	3	AAC59687	Aac59687 Human sec
30	81.2	8.2	351	6	ABL86336	Abl86336 Human ova
31	81.2	8.2	361	6	ABL86348	Abl86348 Human ova
32	81.2	8.2	389	6	ABL87805	Abl87805 Human ova
33	81.2	8.2	573	4	AAK66372	Aak66372 Human imm
34	81	8.2	536	6	ABN61764	Abn61764 Human can
35	81	8.2	4332	4	AAL03068	Aal03068 Human rep
36	81	8.2	4332	4	ABL97404	Abl97404 Human tes
37	81	8.2	4332	7	ABZ68021	Abz68021 Human sec
38	81	8.2	4332	7	ABZ74467	Abz74467 Secreted
39	81	8.2	4332	7	ADA98887	Ada98887 Human sec
40	81	8.2	4332	9	ADA44491	Ada44491 Human sec
41	81	8.2	4332	9	ADC20912	Adc20912 Human sec
42	80.8	8.2	415	4	AAI82427	Aai82427 Human pol
43	80.8	8.2	1793	9	AAD09384	Aad09384 Human ves
44	80.8	8.2	2488	9	ADB63652	Adb63652 Human CDN
45	80.8	8.2	15037	4	AAI99283	Aai99283 Human exc

ALIGNMENTS

RESULT 1
ABL60606
ID   ABL60606 standard; cDNA; 990 BP.
XX
AC   ABL60606;
XX
DT   27-AUG-2002 (first entry)
XX
DE   Human ERCoA3 protein encoding cDNA.
XX
KW   Estrogen Receptor Coregulator 3; ERCoA3; tamoxifen; estrogen; cancer;
KW   osteoporosis; cytostatic; osteopathic; human; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   203..442
FT   /tag= a
FT   /product= "ERCoA3 protein"
XX
PN   WO200228352-A2.
XX
PD   11-APR-2002.
XX
PF   05-OCT-2001; 2001WO-US031271.
XX
PR   05-OCT-2000; 2000US-0238190P.
XX
PA   (UYCA-) UNIV CASE WESTERN RESERVE.
XX
PI   Montano M, Sutton A;
XX
DR   WPI; 2002-454492/48.
DR   P-PSDB; ABS08035.
XX
PT   New polypeptide, that is a functional equivalent of ERCoA3 (Estrogen
PT   Receptor Coregulator 3), is useful in inhibiting or reducing tamoxifen or
PT   estrogen-induced proliferation of cancer cells and in treating
PT   osteoporosis.
XX
PS   Claim 5; Fig 1; 39pp; English.
XX
CC   The invention relates to a ERCoA3 (Estrogen Receptor Coregulator 3)
CC   protein and encoding polynucleotides. ERCoA3 can be used to inhibit or
CC   reduce tamoxifen or estrogen induced proliferation of cancer cells, by
CC   reduced activity of ERCoA3, and for detecting cancer cells that are

CC   tamoxifen resistant, or to treat osteoporosis, by increasing levels of
CC   ERCoA3 in cells. The encoding polynucleotide can be used to inhibit
CC   translation of a mRNA encoding ERCoA3. ERCoA3 acts as a coregulator
CC   protein and can bind to the estrogen receptor to activate a molecular or
CC   cellular response in the cell. The present sequence represents the human
CC   ERCoA3 encoding cDNA
XX
SQ   Sequence 990 BP; 189 A; 281 C; 317 G; 203 T; 0 U; 0 Other;

Query Match   100.0%; Score 990; DB 6; Length 990;
Best Local Similarity 100.0%; Pred. No. 3.7e-235;
Matches 990; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GGAATTGTTCTCGAGGCCAAGAGAATTCGGCAAGGCGATCTCCTGACCTCGTGATCCGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGAATTGTTCTCGAGGCCAAGAGAATTCGGCAAGGCGATCTCCTGACCTCGTGATCCGCCG 60

Qy     61 CCTCAGCTTCCCAAAGTGCTGGGATTACAGGCGATGAGCCACCGACCCGGCCCCGCACAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCTCAGCTTCCCAAAGTGCTGGGATTACAGGCGATGAGCCACCGACCCGGCCCCGCACAC 120

Qy    121 CTAGTTTTAAAGGGCCCCCTGCTGTTGCTGCCGCTGCCTGCCGCCTCCCAGCTGCCCAGTCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTAGTTTTAAAGGGCCCCCTGCTGTTGCTGCCGCTGCCTGCCGCCTCCCAGCTGCCCAGTCTG 180

Qy    181 GCGGGCTCAGTCCCGGTTGCCATGTGTGGGAGACCCGTCCGTAAGCGCTGGATGTGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCGGGCTCAGTCCCGGTTGCCATGTGTGGGAGACCCGTCCGTAAGCGCTGGATGTGG 240

Qy    241 CTTCGCTGATGCACATTGGACACGGGGCTCTGGACTGGGCTAGGGGAAGGGCAGGAGGGGCGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTTCGCTGATGCACATTGGACACGGGGCTCTGGACTGGGCTAGGGGAAGGGCAGGAGGGGCGG 300

Qy    301 AATTGGGCCCGAGGGCCAGGCCTCGCGCACCCCGACCCCCGACCCCGGTGGCCCCGCAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AATTGGGCCCGAGGGCCAGGCCTCGCGCACCCCGACCCCCGACCCCGGTGGCCCCGCAG 360

Qy    361 CGCCTCCCGGTGGCCCTGGAGTGCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGGCGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGCCTCCCGGTGGCCCTGGAGTGCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGGCGA 420

Qy    421 GCTGTGCATGGGCGGTGGCTAAGGCCGTGTTGTTGTTACGATTGGCCAGCGGGGACTTAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCTGTGCATGGGCGGTGGCTAAGGCCGTGTTGTTGTTACGATTGGCCAGCGGGGACTTAAG 480

Qy    481 TGTTGTCTCTGAAGAGCATGGACATTAGTCTGGAGGGTCCTGGAAGAGTGATCCCGCCC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGTTGTCTCTGAAGAGCATGGACATTAGTCTGGAGGGTCCTGGAAGAGTGATCCCGCCC 540

Qy    541 CACCATCAAATGGCGCTTAGTCTTAGGAAGCGGGGTGTGGGTGGGGCCTTAGGGCGGAGGCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CACCATCAAATGGCGCTTAGTCTTAGGAAGCGGGGTGTGGGTGGGGCCTTAGGGCGGAGGCG 600

Qy    601 CAGACATACCCCGAAGTGGTTGGATTGTATACCGCAAGGGGCTGATGCAACCCCCCAAA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CAGACATACCCCGAAGTGGTTGGATTGTATACCGCAAGGGGCTGATGCAACCCCCCAAA 660

Qy    661 GACACTGGAAGGCGTGTGTGGCTGCTGAGGAGGGCCCGCAATCCAGTGTGTCGTGGGCTTTAC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GACACTGGAAGGCGTGTGTGGCTGCTGAGGAGGGCCCGCAATCCAGTGTGTCGTGGGCTTTAC 720

Qy    721 AGGAAGAGCTCCACCTTCTTGAGTGTGCAGATGCGATCTAGGTGTGTCCACCCGATGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AGGAAGAGCTCCACCTTCTTGAGTGTGCAGATGCGATCTAGGTGTGTCCACCCGATGG 780

Qy    781 GAGCTGCGGGCGGCAGATGCTGCCCAGTACAAAGTGATTTGGACCTGGGCCTTCTG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGCTGCGGGCGGCAGATGCTGCCCAGTACAAAGCTGATTTGGACCTGGGCCCTTCTG 840

Qy    841 GACTTCCCTGATTCTCGCTTGCATCTCCAGCAAGTCCTGCCCGTTGGCTGCCTTCAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GACTTCCCTGATTCTCGCTTGCATCTCCAGCAAGTCCTGCCCGTTGGCTGCCTTCAT 900

Qy    901 CCACTCCTCACTTCTGCCTTCAGAGTAAAATTGCAAGATCTCTGGTCGTGCAAAAAAAA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCACTCCTCACTTCTGCCTTCAGAGTAAAATTGCAAGATCTCTGGTCGTGCAAAAAAAA 960

Qy    961 AAAAAAAAAAAAAAAATGAGCGGCCGCAAGTT 990
          ||||||||||||||||||||||||||||||||
Db    961 AAAAAAAAAAAAAAAATGAGCGGCCGCAAGTT 990
```

RESULT 2
AAS69040/c
ID   AAS69040 standard; cDNA; 597 BP.
XX
AC   AAS69040;
XX
DT   13-FEB-2002 (first entry)
XX
DE   DNA encoding novel human diagnostic protein #4844.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C,   Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG04853.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 4844; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have application in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 597 BP; 97 A; 217 C; 196 G; 87 T; 0 U; 0 Other;

Query Match   15.0%;   Score 148.6; DB 5; Length 597;
Best Local Similarity 91.8%;   Pred. No. 6.9e-27;

Matches   157;   Conservative   0;   Mismatches   14;   Indels   0;   Gaps   0;

Qy   128 TAAAGGGCCCTGCGTTGCTGCCCGCTGCCCGCTGCCCAGCTGCCCAGTCTGGCGGGGCT 187
           | |||||||||||| || |||| ||||| |||||||||||||||||||||||||||||
Db   597 TTACAGGCCCCTGCGTTGCTGCGCCGCTGCGCCGCTGCCCAGCTGCCCAGTCTGGCGGGGCT 538

Qy   188 CAGTCCCGGTTGCCATGTGTGGGAGACCGGCGTCGCGGTAAGCGGCTGGATGTGGCTTCGCT 247
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   537 CAGTCCCGGTTGCCATGTGTGGGAGACCGGCGTCGCGGTAAGCGGCTGGATGTGGCTTCGCT 478

Qy   248 GATGCACATTGGACCGGCCTGGACTGGCTAGGGGAAGGGCACGAGGGC 298
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   477 GATGCACATTGGACCGGCCTGGACTGGCTAGGGGAAGGGCACGAGGGCC 427

RESULT 3
AAH18715/c
ID   AAH18715 standard; cDNA; 1602 BP.
XX
AC   AAH18715;
XX
DT   26-JUN-2001 (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:18986.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000;   2000JP-00118776.
PR   02-MAY-2000;   2000JP-00183767.
PR   09-JUN-2000;   2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI   Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 8; SEQ ID NO 18986; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesizing 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesizing polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 1602 BP; 334 A; 458 C; 491 G; 319 T; 0 U; 0 Other;

Query Match   8.8%; Score 87.4; DB 4; Length 1602;
Best Local Similarity   85.8%; Pred. No. 1.4e-11;
Matches   97; Conservative   0; Mismatches   16; Indels   0; Gaps   0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   935 CGATCTCCTGACCTCGTGATCCGCTGATCCAGCCTCTCAAAGTGCTGGGATTACAGGCG 876

Qy   94 TGAGCCACCGCGCCCGGCCCCGACCACCTAGTTTTAAAGGGCCCCTGCTGTTG 146
           |||||||||||||||||||||| |||||||||
Db   875 TGAGCCACCGCGCCCGGCCCCGTCCCCAGGCCCCCTTTTTTTAAGGACCCCCAGTCCTTG 823

RESULT 4
ADC85452
ID   ADC85452 standard; DNA; 67782 BP.
XX
AC   ADC85452;
XX
DT   01-JAN-2004 (first entry)
XX
DE   Human Git2 genomic sequence.
XX
KW   Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW   secreted; transmembrane; intracellular; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003045230-A2.
XX
PD   05-JUN-2003.
XX
PF   02-DEC-2002; 2002WO-US038582.
XX
PR   30-NOV-2001; 2001US-00997722.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW, Engelhard EK;
XX
DR   WPI; 2003-513603/48.
XX
PT   New recombinant nucleic acid comprising a nucleotide sequence of any of
PT   the carcinoma-associated (CA) genes, useful for screening for drug
PT   candidates for diagnosing or treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 238; 983pp; English.
XX
CC   The invention relates to a recombinant nucleic acid comprising a
CC   nucleotide sequence selected from any of the fully defined carcinoma-
CC   associated (CA) genes given from the 50 tables given in the specification. The
CC   CA proteins are secreted, transmembrane or intracellular proteins. The
CC   recombinant nucleic acids are useful for screening for drug candidates
CC   for diagnosing or treating carcinomas. Sequences given in ADC85215-
CC   ADC85514 represent CA genes of the invention.
XX
SQ   Sequence 67782 BP; 18941 A; 14566 C; 14980 G; 18926 T; 0 U; 369 Other;

Query Match   8.5%; Score 84.2; DB 9; Length 67782;
Best Local Similarity   81.0%; Pred. No. 3e-10;
Matches   98; Conservative   0; Mismatches   23; Indels   0; Gaps   0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
           |||||||||||||||||||||||| ||||||||||||||||
Db 51741 CGATCTCCTGACCTCGTGATCCGCCGCCTCGGCCCTGCGGCCCCGGCCCTCCCAAAGTGCTGGGATTACAGGCG 51800

Qy   94 TGAGCCACCGCGCCGCCGGCCCCCGACACCTAGTTTAAAGGGCCCCCTGCTGTTGCTGCCGC 153
                |||||||||||||||||||||||||||||||| ||| || ||||||| ||||||||||||
Db   51801 TGAGCCACCGCGCCGCCGGCCCCCGGCCCAGTGCCTAGTTCTGGCTCCAGT 51860

Qy   154 T 154
                |
Db   51861 T 51861

RESULT 5
ADA02972
ID   ADA02972 standard; DNA; 67783 BP.
XX
AC   ADA02972;
XX
DT   06-NOV-2003 (first entry)
XX
DE   Human GIT2 carcinoma associated gene, SEQ ID NO:1490.
XX
KW   Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW   prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200307146-A2.
XX
PD   17-JUL-2003.
XX
PF   26-DEC-2002; 2002WO-US041414.
XX
PR   26-DEC-2001; 2001US-00035832.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-587068/55.
XX
PT   New recombinant nucleic acid encoding carcinoma associated protein,
PT   useful for preparing compositions for treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 1490; 245pp; English.
XX
CC   The invention relates to recombinant carcinoma associated (CA) nucleic
CC   acid sequences from mouse and human (ADA01482-ADA03094), and to
CC   recombinant carcinoma associated proteins (CAP) encoded by them. The
CC   invention also encompasses expression vectors and host cells comprising a
CC   CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC   binds to the protein, and a biochip comprising CA nucleic acid or
CC   fragments thereof. The sequences of the invention were identified using
CC   oncogenic retroviruses, which insert into the genome of the host organism
CC   at random. Many of these do not carry transduced host oncogenes or
CC   pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC   direct consequence of the effects of proviral integration into host
CC   protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC   carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC   leukaemia) or a propensity to carcinoma by determination of the sequence
CC   of a CA gene, or by determination of CA gene expression in particular
CC   tissues. CA nucleic acids, proteins and antibodies are also useful as
CC   therapeutic agents and in screening and evaluating drug candidates. The
CC   present sequence represents a specifically claimed human CA nucleic acid
CC   sequence of the invention. Note: The complete sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 67783 BP; 18942 A; 14566 C; 14980 G; 18926 T; 0 U; 369 Other;

Query Match   8.5%; Score 84.2; DB 8; Length 67783;
Best Local Similarity 81.0%; Pred. No. 3e-10;
Matches   98; Conservative   0; Mismatches 23; Indels   0; Gaps   0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCCGGCCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
                ||||||||||||||||||||||||||||| ||||||| |||||||||||||||||||||||
Db   51742 CGATCTCCTGACCTCGTGATCCGCCCGGCCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCG 51801

Qy   94 TGAGCCACCGCGCCGCCGGCCCCCGACACCTAGTTTAAAGGGCCCCCTGCTGTTGCTGCCGC 153
                |||||||||||||||||||||||||||||||| ||| || ||||||| ||||||||||||
Db   51802 TGAGCCACCGCGCCGCCGGCCCCCGGCCCAGTGCCTAGTTCTGGCTCCAGT 51861

Qy   154 T 154
                |
Db   51862 T 51862

RESULT 6
ADB72710
ID   ADB72710 standard; DNA; 67783 BP.
XX
AC   ADB72710;
XX
DT   04-DEC-2003 (first entry)
XX
DE   Human GIT2 gene.
XX
KW   human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW   cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS   Homo sapiens.
XX
PN   WO2003008583-A2.
XX
PD   30-JAN-2003.
XX
PF   26-DEC-2001; 2001WO-US051291.
XX
PR   02-MAR-2001; 2001US-00798586.
PR   23-OCT-2001; 2001US-00004113.
PR   08-NOV-2001; 2001US-00052482.
PR   30-NOV-2001; 2001US-00997722.
PR   20-DEC-2001; 2001US-00034650.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW, Engelhard EK;
XX
DR   WPI; 2003-239337/23.
XX
PT   New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT   cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS   Claim 1; SEQ ID NO 538; 2304pp; English.
XX
CC   The invention relates to a novel recombinant nucleic acid comprising a
CC   nucleotide sequence selected from any of the 660 sequences fully defined
CC   in the specification. A polynucleotide of the invention has cytostatic
CC   activity, and may have a use in gene therapy, or in a vaccine. The
CC   recombinant nucleic acids and polypeptides are useful for treating
CC   carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC   sarcomas. The present sequence represents a human gene of the invention.
XX
SQ   Sequence 67783 BP; 18942 A; 14566 C; 14980 G; 18926 T; 0 U; 369 Other;

Query Match   8.5%; Score 84.2; DB 9; Length 67783;
Best Local Similarity 81.0%; Pred. No. 3e-10;
Matches   98; Conservative   0; Mismatches 23; Indels   0; Gaps   0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCCGGCCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
                ||||||||||||||||||||||||||||| ||||||| |||||||||||||||||||||||
Db   51742 CGATCTCCTGACCTCGTGATCCGCCCGGCCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCG 51801

Qy   94 TGAGCCACCGCGCCGCCGGCCCCCGACACCTAGTTTAAAGGGCCCCCTGCTGTTGCTGCCGC 153
                |||||||||||||||||||||||||||||||| ||| || ||||||| ||||||||||||
Db   51802 TGAGCCACCGCGCCGCCGGCCCCGGCCCAGTGCTCTAGTTCTGGCTCCAGT 51861

Qy   154 T 154
            |
Db   51862 T 51862

RESULT 7
AAC19185
ID   AAC19185 standard; cDNA; 274 BP.
XX
AC   AAC19185;
XX
DT   06-OCT-2000 (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 23260.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999; 99US-0122487P.
XX
PA   (GEST) GENSET.
XX
PI   Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 23260; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC   ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC   to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors
XX
SQ   Sequence 274 BP; 50 A; 88 C; 59 G; 77 T; 0 U; 0 Other;

Query Match   8.4%; Score 83.6; DB 3; Length 274;
Best Local Similarity 90.8%; Pred. No. 6.7e-11;
Matches   89; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCAAAGTGCTGGGATTACAGGCA 93
           ||||||||||||| |||||||||||||||||| || ||||||||||||||||||||||
Db   126 CGATCTCCTCACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 185

Qy   94 TGAGGCACCGCGCCCGGCCCCCGACACCTAGTTTTAAA 131
           ||||||||||||||||||||| |||||||||||||||
Db   186 TGAGGCACCGCGCCCGGCCCCCAATCATCTGTTTTAAA 223

RESULT 8
AAH99495/c
ID   AAH99495 standard; cDNA; 2105 BP.

XX
AC   AAH99495;
XX
DT   16-OCT-2001 (first entry)
XX
DE   Human protein encoding cDNA sequence SEQ ID NO:330.
XX
KW   Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW   antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW   antibacterial; endocrine; cardiant; central nervous system; virucide;
KW   anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW   antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW   dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW   neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW   immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW   antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW   cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW   genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW   thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW   allergic rhinitis; diabetes; multiple sclerosis; depression;
KW   Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW   neurological disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153455-A2.
XX
PD   26-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-US035017.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00448725.
PR   25-APR-2000; 2000US-00552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Drmanac RT;
XX
DR   WPI; 2001-457603/49.
DR   P-PSDB; AAM25554.
XX
PT   Isolated human polynucleotides encoding polypeptides, useful for the
PT   treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
PS   Claim 1; Page 444; 1217pp; English.
XX
CC   AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC   AAM25963. The proteins can have activities based on the tissues and cells
CC   they are expressed in, such as: antiinflammatory; antirheumatic;
CC   antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC   central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC   cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC   antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC   antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC   antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC   encoding them can be used in gene therapy, antisense therapy and vaccine
CC   production. The proteins and polynucleotides are useful for screening for
CC   agonists or antagonists of a protein and for the treatment and diagnosis
CC   of disorders associated with the activity of a protein e.g. inflammation,
CC   rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC   neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC   infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC   anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC   osteoporosis, severe combined immunodeficiency, eczema, allergic
CC   rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC   Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC   neurological disorders
XX
SQ   Sequence 2105 BP; 580 A; 460 C; 582 G; 483 T; 0 U; 0 Other;

Query Match   8.4%; Score 83.6; DB 4; Length 2105;
Best Local Similarity 90.8%; Pred. No. 1.3e-10;

```
Matches   89;   Conservative   0;   Mismatches   9;   Indels   0;   Gaps   0;

Qy     34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          ||||||||| |||||| |||||| |||||||| || ||||||||||||||||||||||||
Db    555 CGATCTCCTCACCTCGTGATCCGCCCTCGGCCTCCGGCCTCCCAAAGTGCTGGGATTACAGGCA 496

Qy     94 TGAGCCACCGGCGCCCGGCCCCGACACCTAGTTTTAAA 131
          ||||||||||||||| ||||||       ||||||||
Db    495 TGAGCCACCGGCGCCCGGCCCCAATCATCTGTTTTTAAA 458
```

RESULT 9
AAK66423/C
ID   AAK66423 standard; DNA; 2214 BP.
XX
AC   AAK66423;
XX
DT   06-NOV-2001   (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21235.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214486P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0319968P.
PR   14-SEP-2000; 2000US-0323397P.
PR   14-SEP-2000; 2000US-0323398P.
PR   14-SEP-2000; 2000US-0323399P.
PR   14-SEP-2000; 2000US-0324400P.
PR   14-SEP-2000; 2000US-0324401P.
PR   14-SEP-2000; 2000US-0330063P.
PR   14-SEP-2000; 2000US-0330064P.
PR   14-SEP-2000; 2000US-0330065P.
PR   21-SEP-2000; 2000US-0334223P.
PR   21-SEP-2000; 2000US-0342274P.
PR   25-SEP-2000; 2000US-0334997P.
PR   25-SEP-2000; 2000US-0334998P.
PR   26-SEP-2000; 2000US-0354484P.
PR   27-SEP-2000; 2000US-0358834P.
PR   27-SEP-2000; 2000US-0358836P.
PR   29-SEP-2000; 2000US-0363327P.
PR   29-SEP-2000; 2000US-0363367P.
PR   29-SEP-2000; 2000US-0363368P.
PR   29-SEP-2000; 2000US-0363369P.
PR   29-SEP-2000; 2000US-0363370P.
PR   02-OCT-2000; 2000US-0368802P.
PR   02-OCT-2000; 2000US-0370037P.
PR   02-OCT-2000; 2000US-0370038P.
PR   02-OCT-2000; 2000US-0370039P.
PR   02-OCT-2000; 2000US-0370040P.
PR   13-OCT-2000; 2000US-0399935P.
PR   20-OCT-2000; 2000US-0399937P.
PR   20-OCT-2000; 2000US-0409960P.
PR   20-OCT-2000; 2000US-0412221P.
PR   20-OCT-2000; 2000US-0417785P.
PR   20-OCT-2000; 2000US-0417786P.
PR   20-OCT-2000; 2000US-0417787P.
PR   20-OCT-2000; 2000US-0441808P.
PR   20-OCT-2000; 2000US-0441809P.
PR   20-OCT-2000; 2000US-0441826P.
PR   01-NOV-2000; 2000US-0446417P.
PR   08-NOV-2000; 2000US-0446474P.
PR   08-NOV-2000; 2000US-0446475P.
PR   08-NOV-2000; 2000US-0446476P.
PR   08-NOV-2000; 2000US-0446477P.
PR   08-NOV-2000; 2000US-0446478P.
PR   08-NOV-2000; 2000US-0446523P.
PR   08-NOV-2000; 2000US-0446524P.
PR   08-NOV-2000; 2000US-0446525P.
PR   08-NOV-2000; 2000US-0446526P.
PR   08-NOV-2000; 2000US-0446527P.
PR   08-NOV-2000; 2000US-0446528P.
PR   08-NOV-2000; 2000US-0446532P.
PR   08-NOV-2000; 2000US-0446609P.
PR   08-NOV-2000; 2000US-0446610P.
PR   08-NOV-2000; 2000US-0446611P.
PR   08-NOV-2000; 2000US-0446613P.
PR   17-NOV-2000; 2000US-0449207P.
PR   17-NOV-2000; 2000US-0449208P.
PR   17-NOV-2000; 2000US-0449209P.
PR   17-NOV-2000; 2000US-0449210P.
PR   17-NOV-2000; 2000US-0449211P.
PR   17-NOV-2000; 2000US-0449212P.
PR   17-NOV-2000; 2000US-0449213P.
PR   17-NOV-2000; 2000US-0449214P.
PR   17-NOV-2000; 2000US-0449215P.
PR   17-NOV-2000; 2000US-0449216P.
```

```
PR   17-NOV-2000;  2000US-0249217P.
PR   17-NOV-2000;  2000US-0249218P.
PR   17-NOV-2000;  2000US-0249244P.
PR   17-NOV-2000;  2000US-0249245P.
PR   17-NOV-2000;  2000US-0249264P.
PR   17-NOV-2000;  2000US-0249265P.
PR   17-NOV-2000;  2000US-0249297P.
PR   17-NOV-2000;  2000US-0249299P.
PR   17-NOV-2000;  2000US-0249300P.
PR   01-DEC-2000;  2000US-0250160P.
PR   01-DEC-2000;  2000US-0250391P.
PR   05-DEC-2000;  2000US-0251030P.
PR   05-DEC-2000;  2000US-0251988P.
PR   06-DEC-2000;  2000US-0356719P.
PR   08-DEC-2000;  2000US-0251479P.
PR   08-DEC-2000;  2000US-0251856P.
PR   08-DEC-2000;  2000US-0251868P.
PR   08-DEC-2000;  2000US-0251869P.
PR   08-DEC-2000;  2000US-0251989P.
PR   08-DEC-2000;  2000US-0251990P.
PR   11-DEC-2000;  2000US-0254097P.
PR   05-JAN-2001;  2001US-0296678P.
XX
PA   (HUMA-)  HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI;  2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure;  SEQ ID NO 21235;  3071pp + Sequence Listing;  English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 2214 BP;  617 A;  455 C;  619 G;  523 T;  0 U;  0 Other;

     Query Match      8.4%;  Score 83.2;  DB 4;  Length 2214;
     Best Local Similarity   83.9%;  Pred. No. 1.7e-10;
     Matches   94;  Conservative   0;  Mismatches   18;  Indels   0;  Gaps   0;

Qy       34  CGATCTCCTGACCTCGTGATCCGCCCCGCCTCCAGCTTCCCAAAGTGCTGGGATTACAGGCA   93
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db      403  CGATCTCCTGACCTCGTGATCCGCCCGTCTCGGCCTCCCCAAAGTGCTGGGACTGCAGGCG  344

Qy       94  TGAGCCACCGCGCCCCCGGCCCCCGACACCTAGTTTTAAAGGGCCCCTGCTGTT  145
             ||||||||||||||| |||| |||||| ||||||||  ||||| ||||||||||
Db      343  TGAGCCACCGCGGCCCGGCCCCAACGTCATTTTCTTTTCAGGAGTCCCTTCTGTT  292
```

```
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21236.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001;  2001WO-US001354.
XX
PR   31-JAN-2000;  2000US-0179065P.
PR   04-FEB-2000;  2000US-0180628P.
PR   24-FEB-2000;  2000US-0184664P.
PR   02-MAR-2000;  2000US-0186350P.
PR   16-MAR-2000;  2000US-0189874P.
PR   17-MAR-2000;  2000US-0190076P.
PR   18-APR-2000;  2000US-0198123P.
PR   19-MAY-2000;  2000US-0205515P.
PR   07-JUN-2000;  2000US-0209467P.
PR   28-JUN-2000;  2000US-0214886P.
PR   30-JUN-2000;  2000US-0215135P.
PR   07-JUL-2000;  2000US-0216647P.
PR   07-JUL-2000;  2000US-0216680P.
PR   11-JUL-2000;  2000US-0217487P.
PR   11-JUL-2000;  2000US-0217496P.
PR   14-JUL-2000;  2000US-0218290P.
PR   26-JUL-2000;  2000US-0220963P.
PR   26-JUL-2000;  2000US-0220964P.
PR   14-AUG-2000;  2000US-0224518P.
PR   14-AUG-2000;  2000US-0224519P.
PR   14-AUG-2000;  2000US-0225213P.
PR   14-AUG-2000;  2000US-0225214P.
PR   14-AUG-2000;  2000US-0225266P.
PR   14-AUG-2000;  2000US-0225267P.
PR   14-AUG-2000;  2000US-0225268P.
PR   14-AUG-2000;  2000US-0225270P.
PR   14-AUG-2000;  2000US-0225447P.
PR   14-AUG-2000;  2000US-0225757P.
PR   14-AUG-2000;  2000US-0225758P.
PR   14-AUG-2000;  2000US-0225759P.
PR   18-AUG-2000;  2000US-0226279P.
PR   22-AUG-2000;  2000US-0226681P.
PR   22-AUG-2000;  2000US-0226868P.
PR   22-AUG-2000;  2000US-0227182P.
PR   23-AUG-2000;  2000US-0227009P.
PR   30-AUG-2000;  2000US-0229924P.
PR   01-SEP-2000;  2000US-0229287P.
PR   01-SEP-2000;  2000US-0229343P.
PR   01-SEP-2000;  2000US-0229344P.
PR   01-SEP-2000;  2000US-0229345P.
PR   05-SEP-2000;  2000US-0229509P.
PR   05-SEP-2000;  2000US-0229513P.
PR   06-SEP-2000;  2000US-0230437P.
PR   06-SEP-2000;  2000US-0230438P.
PR   08-SEP-2000;  2000US-0231242P.
PR   08-SEP-2000;  2000US-0231243P.
PR   08-SEP-2000;  2000US-0231244P.
PR   08-SEP-2000;  2000US-0231413P.
PR   08-SEP-2000;  2000US-0231414P.
PR   08-SEP-2000;  2000US-0232080P.
PR   08-SEP-2000;  2000US-0232081P.
PR   12-SEP-2000;  2000US-0231968P.
PR   14-SEP-2000;  2000US-0232397P.
PR   14-SEP-2000;  2000US-0232398P.
PR   14-SEP-2000;  2000US-0232399P.
PR   14-SEP-2000;  2000US-0232400P.
PR   14-SEP-2000;  2000US-0232401P.
PR   14-SEP-2000;  2000US-0233063P.
```

```
RESULT 10
AAK66424/c
ID   AAK66424 standard; DNA; 6742 BP.
XX
AC   AAK66424;
XX
```

```
PR   14-SEP-2000;  2000US-0233064P.
PR   14-SEP-2000;  2000US-0233065P.
PR   21-SEP-2000;  2000US-0234223P.
PR   21-SEP-2000;  2000US-0234274P.
PR   25-SEP-2000;  2000US-0234997P.
PR   25-SEP-2000;  2000US-0234998P.
PR   26-SEP-2000;  2000US-0235484P.
PR   27-SEP-2000;  2000US-0235834P.
PR   27-SEP-2000;  2000US-0235836P.
PR   29-SEP-2000;  2000US-0236327P.
PR   29-SEP-2000;  2000US-0236367P.
PR   29-SEP-2000;  2000US-0236368P.
PR   29-SEP-2000;  2000US-0236369P.
PR   29-SEP-2000;  2000US-0236370P.
PR   02-OCT-2000;  2000US-0236802P.
PR   02-OCT-2000;  2000US-0237037P.
PR   02-OCT-2000;  2000US-0237038P.
PR   02-OCT-2000;  2000US-0237039P.
PR   02-OCT-2000;  2000US-0237040P.
PR   13-OCT-2000;  2000US-0239935P.
PR   13-OCT-2000;  2000US-0239937P.
PR   20-OCT-2000;  2000US-0240060P.
PR   20-OCT-2000;  2000US-0241221P.
PR   20-OCT-2000;  2000US-0241785P.
PR   20-OCT-2000;  2000US-0241786P.
PR   20-OCT-2000;  2000US-0241787P.
PR   20-OCT-2000;  2000US-0241808P.
PR   20-OCT-2000;  2000US-0241809P.
PR   20-OCT-2000;  2000US-0241826P.
PR   01-NOV-2000;  2000US-0244617P.
PR   01-NOV-2000;  2000US-0244674P.
PR   08-NOV-2000;  2000US-0246475P.
PR   08-NOV-2000;  2000US-0246476P.
PR   08-NOV-2000;  2000US-0246477P.
PR   08-NOV-2000;  2000US-0246478P.
PR   08-NOV-2000;  2000US-0246523P.
PR   08-NOV-2000;  2000US-0246524P.
PR   08-NOV-2000;  2000US-0246525P.
PR   08-NOV-2000;  2000US-0246526P.
PR   08-NOV-2000;  2000US-0246527P.
PR   08-NOV-2000;  2000US-0246528P.
PR   08-NOV-2000;  2000US-0246532P.
PR   08-NOV-2000;  2000US-0246609P.
PR   08-NOV-2000;  2000US-0246610P.
PR   08-NOV-2000;  2000US-0246611P.
PR   08-NOV-2000;  2000US-0246613P.
PR   17-NOV-2000;  2000US-0249207P.
PR   17-NOV-2000;  2000US-0249208P.
PR   17-NOV-2000;  2000US-0249209P.
PR   17-NOV-2000;  2000US-0249210P.
PR   17-NOV-2000;  2000US-0249211P.
PR   17-NOV-2000;  2000US-0249212P.
PR   17-NOV-2000;  2000US-0249213P.
PR   17-NOV-2000;  2000US-0249214P.
PR   17-NOV-2000;  2000US-0249215P.
PR   17-NOV-2000;  2000US-0249216P.
PR   17-NOV-2000;  2000US-0249217P.
PR   17-NOV-2000;  2000US-0249218P.
PR   17-NOV-2000;  2000US-0249244P.
PR   17-NOV-2000;  2000US-0249245P.
PR   17-NOV-2000;  2000US-0249264P.
PR   17-NOV-2000;  2000US-0249265P.
PR   17-NOV-2000;  2000US-0249297P.
PR   17-NOV-2000;  2000US-0249299P.
PR   17-NOV-2000;  2000US-0249300P.
PR   01-DEC-2000;  2000US-0250160P.
PR   01-DEC-2000;  2000US-0250391P.
PR   05-DEC-2000;  2000US-0251030P.
PR   05-DEC-2000;  2000US-0251988P.
PR   05-DEC-2000;  2000US-0256719P.
PR   06-DEC-2000;  2000US-0251479P.
PR   06-DEC-2000;  2000US-0251856P.
PR   08-DEC-2000;  2000US-0251868P.
```

```
PR   08-DEC-2000;  2000US-0251869P.
PR   08-DEC-2000;  2000US-0251989P.
PR   08-DEC-2000;  2000US-0251990P.
PR   11-DEC-2000;  2000US-0254097P.
PR   05-JAN-2001;  2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure; SEQ ID NO 21236; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 6742 BP; 2027 A; 1259 C; 1693 G; 1762 T; 0 U; 1 Other;

Query Match             8.4%;  Score 83.2;  DB 4;  Length 6742;
Best Local Similarity   83.9%;  Pred. No. 2.4e-10;
Matches   94;  Conservative   0;  Mismatches  18;  Indels   0;  Gaps   0;

Qy     34 CGATCTCCTGACCTCTGTGATCCGCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA  93
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   4920 CGATCTCCTGACCTCTGTGATCCGCCGTCTCCGGCCTCCCAAAGTGCTGGGACTGCAGGCG 4861

Qy     94 TGAGCCACCGCGCCCGGCCCCGACACCTAGTTTTAAAGGGCCCCTGCTGTT 145
          ||||||||||||||||||||||||||||| || ||||| |||||||||||
Db   4860 TGAGCCACCGCGCCCGGCCCCGCCAACGTCATTTTCTTTTCAGGAGTCCCTTCTGTT 4809


RESULT 11
AAK66425/c
ID   AAK66425 standard; DNA; 10772 BP.
XX
AC   AAK66425;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21237.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
```

```
PR  24-FEB-2000;  2000US-0184664P.        PR  02-OCT-2000;  2000US-0237039P.
PR  02-MAR-2000;  2000US-0186350P.        PR  02-OCT-2000;  2000US-0237040P.
PR  16-MAR-2000;  2000US-0189874P.        PR  13-OCT-2000;  2000US-0239935P.
PR  17-MAR-2000;  2000US-0190076P.        PR  13-OCT-2000;  2000US-0239937P.
PR  18-APR-2000;  2000US-0198123P.        PR  20-OCT-2000;  2000US-0240960P.
PR  19-MAY-2000;  2000US-0205515P.        PR  20-OCT-2000;  2000US-0241221P.
PR  07-JUN-2000;  2000US-0209467P.        PR  20-OCT-2000;  2000US-0241785P.
PR  28-JUN-2000;  2000US-0214886P.        PR  20-OCT-2000;  2000US-0241786P.
PR  30-JUN-2000;  2000US-0215135P.        PR  20-OCT-2000;  2000US-0241787P.
PR  07-JUL-2000;  2000US-0216647P.        PR  20-OCT-2000;  2000US-0241808P.
PR  07-JUL-2000;  2000US-0216880P.        PR  20-OCT-2000;  2000US-0241809P.
PR  11-JUL-2000;  2000US-0217487P.        PR  20-OCT-2000;  2000US-0241826P.
PR  11-JUL-2000;  2000US-0174496P.        PR  01-NOV-2000;  2000US-0244617P.
PR  14-JUL-2000;  2000US-0218290P.        PR  08-NOV-2000;  2000US-0246474P.
PR  26-JUL-2000;  2000US-0209963P.        PR  08-NOV-2000;  2000US-0246475P.
PR  26-JUL-2000;  2000US-0209964P.        PR  08-NOV-2000;  2000US-0246476P.
PR  14-AUG-2000;  2000US-0224518P.        PR  08-NOV-2000;  2000US-0246477P.
PR  14-AUG-2000;  2000US-0224519P.        PR  08-NOV-2000;  2000US-0246478P.
PR  14-AUG-2000;  2000US-0225213P.        PR  08-NOV-2000;  2000US-0246523P.
PR  14-AUG-2000;  2000US-0225214P.        PR  08-NOV-2000;  2000US-0246524P.
PR  14-AUG-2000;  2000US-0225266P.        PR  08-NOV-2000;  2000US-0246525P.
PR  14-AUG-2000;  2000US-0225267P.        PR  08-NOV-2000;  2000US-0246526P.
PR  14-AUG-2000;  2000US-0225268P.        PR  08-NOV-2000;  2000US-0246527P.
PR  14-AUG-2000;  2000US-0225270P.        PR  08-NOV-2000;  2000US-0246528P.
PR  14-AUG-2000;  2000US-0225447P.        PR  08-NOV-2000;  2000US-0246532P.
PR  14-AUG-2000;  2000US-0225757P.        PR  08-NOV-2000;  2000US-0246609P.
PR  14-AUG-2000;  2000US-0225758P.        PR  08-NOV-2000;  2000US-0246610P.
PR  14-AUG-2000;  2000US-0225759P.        PR  08-NOV-2000;  2000US-0246611P.
PR  18-AUG-2000;  2000US-0226279P.        PR  08-NOV-2000;  2000US-0246613P.
PR  22-AUG-2000;  2000US-0226681P.        PR  17-NOV-2000;  2000US-0249207P.
PR  22-AUG-2000;  2000US-0226868P.        PR  17-NOV-2000;  2000US-0249208P.
PR  22-AUG-2000;  2000US-0227182P.        PR  17-NOV-2000;  2000US-0249209P.
PR  23-AUG-2000;  2000US-0227009P.        PR  17-NOV-2000;  2000US-0249210P.
PR  30-AUG-2000;  2000US-0228924P.        PR  17-NOV-2000;  2000US-0249211P.
PR  01-SEP-2000;  2000US-0229287P.        PR  17-NOV-2000;  2000US-0249212P.
PR  01-SEP-2000;  2000US-0229343P.        PR  17-NOV-2000;  2000US-0249213P.
PR  01-SEP-2000;  2000US-0229344P.        PR  17-NOV-2000;  2000US-0249214P.
PR  01-SEP-2000;  2000US-0229345P.        PR  17-NOV-2000;  2000US-0249215P.
PR  05-SEP-2000;  2000US-0229509P.        PR  17-NOV-2000;  2000US-0249216P.
PR  05-SEP-2000;  2000US-0229513P.        PR  17-NOV-2000;  2000US-0249217P.
PR  06-SEP-2000;  2000US-0230437P.        PR  17-NOV-2000;  2000US-0249218P.
PR  08-SEP-2000;  2000US-0230438P.        PR  17-NOV-2000;  2000US-0249244P.
PR  08-SEP-2000;  2000US-0231242P.        PR  17-NOV-2000;  2000US-0249245P.
PR  08-SEP-2000;  2000US-0231243P.        PR  17-NOV-2000;  2000US-0249264P.
PR  08-SEP-2000;  2000US-0231244P.        PR  17-NOV-2000;  2000US-0249265P.
PR  08-SEP-2000;  2000US-0231413P.        PR  17-NOV-2000;  2000US-0249297P.
PR  08-SEP-2000;  2000US-0231414P.        PR  17-NOV-2000;  2000US-0249299P.
PR  08-SEP-2000;  2000US-0232080P.        PR  17-NOV-2000;  2000US-0249300P.
PR  08-SEP-2000;  2000US-0232081P.        PR  01-DEC-2000;  2000US-0250160P.
PR  12-SEP-2000;  2000US-0231968P.        PR  05-DEC-2000;  2000US-0250391P.
PR  14-SEP-2000;  2000US-0232397P.        PR  05-DEC-2000;  2000US-0251030P.
PR  14-SEP-2000;  2000US-0232398P.        PR  05-DEC-2000;  2000US-0251988P.
PR  14-SEP-2000;  2000US-0232399P.        PR  06-DEC-2000;  2000US-0256719P.
PR  14-SEP-2000;  2000US-0232400P.        PR  08-DEC-2000;  2000US-0251479P.
PR  14-SEP-2000;  2000US-0232401P.        PR  08-DEC-2000;  2000US-0251856P.
PR  14-SEP-2000;  2000US-0233063P.        PR  08-DEC-2000;  2000US-0251868P.
PR  14-SEP-2000;  2000US-0233064P.        PR  08-DEC-2000;  2000US-0251869P.
PR  14-SEP-2000;  2000US-0233065P.        PR  08-DEC-2000;  2000US-0251989P.
PR  21-SEP-2000;  2000US-0234223P.        PR  08-DEC-2000;  2000US-0251990P.
PR  21-SEP-2000;  2000US-0234274P.        PR  11-DEC-2000;  2000US-0254097P.
PR  21-SEP-2000;  2000US-0234997P.        PR  05-JAN-2001;  2001US-0259678P.
PR  25-SEP-2000;  2000US-0234998P.        XX
PR  26-SEP-2000;  2000US-0235484P.        PA  (HUMA-) HUMAN GENOME SCI INC.
PR  27-SEP-2000;  2000US-0235834P.        XX
PR  27-SEP-2000;  2000US-0235836P.        PI  Rosen CA, Barash SC, Ruben SM;
PR  29-SEP-2000;  2000US-0236327P.        XX
PR  29-SEP-2000;  2000US-0236367P.        DR  WPI; 2001-483426/52.
PR  29-SEP-2000;  2000US-0236368P.        XX
PR  29-SEP-2000;  2000US-0236369P.        PT  Nucleic acids encoding human immune/hematopoietic antigen polyepptides,
PR  29-SEP-2000;  2000US-0236370P.        PT  useful for preventing, diagnosing and/or treating cancers and metastasis.
PR  02-OCT-2000;  2000US-0236802P.        XX
PR  02-OCT-2000;  2000US-0237037P.        PS  Disclosure; SEQ ID NO 21237; 3071pp + Sequence Listing; English.
PR  02-OCT-2000;  2000US-0237038P.        XX
```

CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 10772 BP; 3111 A; 2137 C; 2495 G; 3029 T; 0 U; 0 Other;

Query Match 8.4%; Score 83.2; DB 4; Length 10772;
Best Local Similarity 83.9%; Pred. No. 2.9e-10;
Matches 94; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          |||||||||||||||||||||||||| ||||||| ||||||||||||||||| ||||||||
Db 8950 CGATCTCCTGACCTCGTGATCCGCCGCCCGTCTCGGCCTCCCAAAGTGCTGGGACTGCAGGCG 8891

Qy   94 TGAGGCACCGCGCCCGGCCCCCGACACCTAGTTTTAAAGGGGCCCCTGCTGTT 145
          ||||||||||||||||||| ||||||||||| ||||| ||| |||||||
Db 8890 TGAGGCACCGCGCCCGGCCCCGGCCAACGTCATTTCATTTCTTTCAGGAGTCCCTTCTGTT 8839

RESULT 12
ADC86452
ID ADC86452 standard; DNA; 13695 BP.
XX
AC ADC86452;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human GPCR gene SEQ ID NO.905.
XX
KW ds; gene; human; GPCR;
KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1270724-A2.
XX
PD 02-JAN-2003.
XX
PF 18-JUN-2002; 2002EP-00013517.
PR 18-JUN-2001; 2001JP-00246789.
XX
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI Suwa M, Asai K, Akiyama Y, Aburatani H;
XX
DR WPI; 2003-315783/31.
DR P-PSDB; ADC86453.
XX
PT New polynucleotide, useful for preparing a composition for treating a
PT patient in need of increased or suppressed activity or expression of the
PT guanosine triphosphate-binding protein coupled receptor.
XX
PS claim 1; SEQ ID NO 905; 28pp; English.
XX
CC The invention relates to a novel polynucleotide encoding a guanosine
CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC the invention may have a use in gene therapy. The polynucleotide and

CC polypeptide are useful for preparing a composition for treating a patient
CC in need of increased or suppressed activity or expression of the
CC guanosine triphosphate-binding protein coupled receptor. The
CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC invention.
XX
SQ Sequence 13695 BP; 2885 A; 3811 C; 3868 G; 3131 T; 0 U; 0 Other;

Query Match 8.4%; Score 83; DB 9; Length 13695;
Best Local Similarity 89.9%; Pred. No. 3.5e-10;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
         ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db 842 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCG 901

Qy   94 TGAGGCACCGCGCCCGGCCCCCGACACCTAGTTTTAAAG 132
         ||||||||||||||||||| ||| ||||
Db 902 TGAGGCACCGCGCCCGGCCTCCAGTTTCCTCTTTTATAG 940

RESULT 13
AAH17556/c
ID AAH17556 standard; cDNA; 2251 BP.
XX
AC AAH17556;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:17037.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 17037; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the

```
CC    specification. The primer sets can be used in antisense therapy and in
CC    gene therapy. The primers are useful for synthesising polynucleotides,
CC    particularly full-length cDNAs. The primers are also useful for the
CC    detection and/or diagnosis of the abnormality of the proteins encoded by
CC    the full-length cDNAs. The primers allow obtaining of the full-length
CC    cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC    AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC    represent human amino acid sequences; AAB92446 to AAH13629 represent
CC    oligonucleotides, all of which are used in the exemplification of the
CC    present invention
XX
SQ    Sequence 2251 BP; 678 A; 479 C; 550 G; 544 T; 0 U; 0 Other;

Query Match             8.3%;  Score 82.6;  DB 4;  Length 2251;
Best Local Similarity   73.1%;  Pred. No. 2.4e-10;
Matches  106;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy    34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          |||||||||||| ||||||| ||| |||||||||||||||||||||||||||||||||||
Db  1740 CGAACCTCCAGACCTCGTGATCCGCCCACCTTGCCTCCCAAAGTGCTGGGATTACAGGCA 1681

Qy    94 TGAGCCACCGCGCCCGGCGCCCCGACACCTAGTTTTAAAGGGCCCCTGCTGTTGCTGCCGC 153
          ||||||||| |||||  || ||| ||| ||||||||  |||||||||||| |||||||||
Db  1680 TGAGCCACCACGCCCAGCTGGCATCTCTATTTTAATGTACCCCAGGTGATTCTGACAC 1621

Qy   154 TGCCGCCGGCTCCCAGCTGCCCAGTC 178
          |||  |||||||||||||||||||||
Db  1620 TGAGCCAGGGCTGAGAACCACTGCC 1596


RESULT 14
AAK85393/c
ID   AAK85393 standard; DNA; 444 BP.
XX
AC   AAK85393;
XX
DT   07-NOV-2001 (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40205.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218200P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225256P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226661P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   21-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0246417P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
```

Query Match   8.3%;   Score 82.2;   DB 4;   Length 444;
Best Local Similarity   87.4%;   Pred. No. 1.7e-10;
Matches 90;   Conservative   0;   Mismatches 13;   Indels   0;   Gaps   0;

Qy   34 CGATCTCTGACCTCGTGATCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          ||||||||||||||||||||| ||||||||||||||||||||||||| |||||||||
Db   223 CGATCTCCTGACCTCGTGATCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 164

Qy   94 TGAGCCACCGCGCCCGGCCCCGACCACCTAGTTTAAAGGGCC 136
          |||||||||||| ||||| |||| ||||| |||||| |||||
Db   163 TGAGCCACCGCGCCCGGCCCGGACTAAAGGAAATTTTAAAAGGGC 121

RESULT 15
ADB63236/c
ID   ADB63236 standard; cDNA; 2390 BP.
XX
AC   ADB63236;
XX
DT   04-DEC-2003 (first entry)
XX
DE   Human cDNA encoding clone SPLEN20204670.
XX
KW   Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
KW   tissue regeneration; cell regeneration; membrane protein;
KW   signal transduction-related protein; transcription-related protein;
KW   osteoporosis; neurological disease; cancer; tumour.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   6..311
FT   /*tag= a
FT   /product= "Clone SPLEN20204670 protein"
XX
PN   EP1308459-A2.
XX
PD   07-MAY-2003.
XX
PF   28-MAR-2002; 2002BP-00007401.
XX
PR   05-NOV-2001; 2001JP-00379298.
PR   25-JAN-2002; 2002US-00350978.
XX
PA   (HELI-) HELIX RES INST.
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI   Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI   Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR   WPI; 2003-450961/43.
DR   P-PSDB; ADB65206.
XX
PT   New polynucleotides and polypeptides, useful for developing a diagnostic
PT   marker or medicines for regulation of their expression and activity, or
PT   as targets of gene therapy.
XX
PS   Claim 1; Page; 222pp; English.
XX
CC   The invention discloses a polynucleotide comprising a sequence selected
CC   from 1970 fully defined nucleotide sequences which encode novel
CC   polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC   or its partial peptide, an antibody binding to the polypeptide or peptide
CC   of the polynucleotide, immunologically assaying the polypeptide or
CC   peptide of the polynucleotide by contacting the polypeptide or peptide
CC   with the antibody of the encoded protein, and observing the binding
CC   between the two, a transformant carrying the polynucleotide in an
CC   expressible manner and an antisense polynucleotide. The oligonucleotide
CC   is useful as a primer for synthesising the polynucleotide, or as a probe
CC   for detecting the polynucleotide. The polynucleotides and encoded
CC   proteins are useful as pharmaceutical agents and many disease-related
CC   genes may be included in them, for developing a diagnostic marker or

PR   08-NOV-2000;   2000US-0246525P.
PR   08-NOV-2000;   2000US-0246526P.
PR   08-NOV-2000;   2000US-0246527P.
PR   08-NOV-2000;   2000US-0246528P.
PR   08-NOV-2000;   2000US-0246532P.
PR   08-NOV-2000;   2000US-0246609P.
PR   08-NOV-2000;   2000US-0246610P.
PR   08-NOV-2000;   2000US-0246611P.
PR   08-NOV-2000;   2000US-0246613P.
PR   17-NOV-2000;   2000US-0249207P.
PR   17-NOV-2000;   2000US-0249208P.
PR   17-NOV-2000;   2000US-0249209P.
PR   17-NOV-2000;   2000US-0249210P.
PR   17-NOV-2000;   2000US-0249211P.
PR   17-NOV-2000;   2000US-0249212P.
PR   17-NOV-2000;   2000US-0249213P.
PR   17-NOV-2000;   2000US-0249214P.
PR   17-NOV-2000;   2000US-0249215P.
PR   17-NOV-2000;   2000US-0249216P.
PR   17-NOV-2000;   2000US-0249217P.
PR   17-NOV-2000;   2000US-0249218P.
PR   17-NOV-2000;   2000US-0249244P.
PR   17-NOV-2000;   2000US-0249245P.
PR   17-NOV-2000;   2000US-0249264P.
PR   17-NOV-2000;   2000US-0249265P.
PR   17-NOV-2000;   2000US-0249297P.
PR   17-NOV-2000;   2000US-0249299P.
PR   17-NOV-2000;   2000US-0249300P.
PR   01-DEC-2000;   2000US-0250160P.
PR   01-DEC-2000;   2000US-0250351P.
PR   05-DEC-2000;   2000US-0251030P.
PR   05-DEC-2000;   2000US-0251988P.
PR   05-DEC-2000;   2000US-0251989P.
PR   06-DEC-2000;   2000US-0256719P.
PR   06-DEC-2000;   2000US-0251479P.
PR   08-DEC-2000;   2000US-0251856P.
PR   08-DEC-2000;   2000US-0251868P.
PR   08-DEC-2000;   2000US-0251869P.
PR   08-DEC-2000;   2000US-0251989P.
PR   08-DEC-2000;   2000US-0251990P.
PR   11-DEC-2000;   2000US-0254097P.
PR   05-JAN-2001;   2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure; SEQ ID NO 40205; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 444 BP; 122 A; 111 C; 126 G; 85 T; 0 U; 0 Other;

CC   medicines for regulation of their expression and activity, or as targets
CC   of gene therapy. The genes are involved in tissue and/or cell
CC   regeneration. Membrane proteins, signal transduction-related proteins,
CC   transcription-related proteins, disease-related proteins and genes
CC   encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC   neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC   the activity or expression of the encoded protein to treat diseases. The
CC   sequence presented is a cDNA of the invention. Note: Some of the sequence
CC   data for this patent is not represented in the printed specification, but
CC   is based on sequence information supplied by the European Patent Office.
XX
SQ   Sequence 2390 BP; 663 A; 460 C; 500 G; 767 T; 0 U; 0 Other;

Query Match
Best Local Similarity   8.3%;   Score 82.2; DB 9;   Length 2390;
                        87.4%;   Pred. No. 3.1e-10;
Matches   90; Conservative   0; Mismatches   13;   Indels   0;   Gaps   0;

Qy   34 CGATCTCCTGACCTGCTGTGATCCGCCCGCCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
           ||||||||||||||| || |||||||||||||| | |||||||||||||||||||||||||||
Db   2200 CGATCTCCTGACCTGCTGTGATCCGCCCGCCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 2141

Qy   94 TGAGCCACCGCGCCCGGCCCCGACACCTAGTTTTAAAGGGCC 136
           ||||||||||||||||||||||| ||| | |||||||||
Db   2140 TGAGCCACCGCGCCCGGCCCGACTAAAGGAAATTTTAAAAGGGC 2098

Search completed: July 10, 2004, 09:54:39
Job time : 502 secs

This Page Blank (uspto)

```
        GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         July 10, 2004, 09:37:04 ; Search time 100 seconds
                                (without alignments)
                        5494.014 Million cell updates/sec

Title:          US-09-972-032-1
Perfect score:  990
Sequence:       1 ggaattgttctcgaggccaa..........aaaatgagcggcgcaagtt 990

Scoring table:  IDENTITY.NUC
                Gapop 10.0 , Gapext 1.0

Searched:       682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:      1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Issued Patents NA:*
                1: /cgn2_6/ptcdata/2/ina/5A_COMB.seq:*
                2: /cgn2_6/ptcdata/2/ina/5B_COMB.seq:*
                3: /cgn2_6/ptcdata/2/ina/6A_COMB.seq:*
                4: /cgn2_6/ptcdata/2/ina/6B_COMB.seq:*
                5: /cgn2_6/ptcdata/2/ina/PCTUS_COMB.seq:*
                6: /cgn2_6/ptcdata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	83.6	8.4	506	4	US-09-621-976-15373	Sequence 15373, A
2	80.4	8.1	51719	4	US-09-918-686-2	Sequence 1, Appli
3	80.4	8.1	92139	4	US-09-918-686-1	Sequence 1, Appli
4	78.8	8.0	511	4	US-09-621-976-9083	Sequence 9083, Ap
5	78.4	7.9	480	4	US-09-621-976-8284	Sequence 8284, Ap
6	78.4	7.9	21721	4	US-09-269-939A-41	Sequence 41, Appl
7	78.4	7.9	22976	4	US-09-269-939A-19	Sequence 19, Appl
8	78.4	7.9	23187	4	US-09-621-976-12615	Sequence 12615, A
9	77.8	7.9	397	4	US-09-499-522-1	Sequence 1, Appli
10	77.6	7.8	2280	4	US-09-702-705-321	Sequence 321, App
11	77.6	7.8	2280	4	US-09-736-457-321	Sequence 321, App
12	77.6	7.8	2280	4	US-09-614-124B-321	Sequence 321, App
13	77.6	7.8	2280	4	US-09-671-325-321	Sequence 321, App
14	77.6	7.8	2280	4	US-09-589-184-321	Sequence 321, App
15	77.6	7.8	319608	4	US-09-539-333D-1	Sequence 1, Appli
16	77.6	7.8	319608	4	US-09-679-409-1	Sequence 1, Appli
17	77.2	7.8	162450	4	US-09-345-882-1	Sequence 5, Appli
18	77	7.8	320	1	US-08-629-939-5	Sequence 5, Appli
19	77	7.8	320	1	US-08-759-873-5	Sequence 5, Appli
20	76.8	7.8	1000	4	US-09-671-317-469	Sequence 469, App
21	76.8	7.8	1000	4	US-09-671-317-484	Sequence 484, App
22	76.8	7.8	49312	4	US-09-671-317-485	Sequence 485, App
23	76.8	7.8	786431	4	US-09-751-389-3	Sequence 3, Appli
24	76.4	7.7	40000	4	US-09-750-049-18	Sequence 18, Appl
25	76.4	7.7	63588	4	US-09-873-404-3	Sequence 3, Appli
26	76.2	7.7	578	4	US-09-495-050A-99	Sequence 99, Appl
27	76.2	7.7	87350	3	US-08-781-891-79	Sequence 79, Appl
28	76.2	7.7	87350	4	US-09-618-166-79	Sequence 79, Appl
29	76.2	7.7	87543	4	US-09-791-211-3	Sequence 3, Appli
30	76.2	7.7	392000	4	US-10-027-983-11	Sequence 11, Appl
31	76	7.7	193303	4	US-09-497-855A-37	Sequence 37, Appl
32	76	7.7	193303	4	US-09-497-855A-44	Sequence 44, Appl
33	75.8	7.7	10380	3	US-09-077-354B-3	Sequence 3, Appli
34	75.6	7.6	480	4	US-09-621-976-12757	Sequence 12757, A
35	75.6	7.6	1001	4	US-09-671-317-223	Sequence 223, App
36	75.6	7.6	246240	2	US-08-724-394A-20	Sequence 20, Appl
37	75.6	7.6	246240	2	US-08-724-394A-21	Sequence 21, Appl
38	75.6	7.6	246240	2	US-08-724-394A-22	Sequence 22, Appl
39	75.4	7.6	40328	3	US-08-742-185-102	Sequence 102, App
40	75.4	7.6	43795	3	US-08-742-185-101	Sequence 101, App
41	75.4	7.6	116592	4	US-09-818-512-3	Sequence 3, Appli
42	75.2	7.6	489	4	US-09-621-976-652	Sequence 652, App
43	75.2	7.6	18073	3	US-09-078-294-12	Sequence 12, Appl
44	75.2	7.6	31208	4	US-09-852-067-3	Sequence 3, Appli
45	75.2	7.6	81001	4	US-09-750-580-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-09-621-976-15373
; Sequence 15373, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT: Jobert, S.
;  APPLICANT: Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE: 2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 15373
;  LENGTH: 506
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-621-976-15373

  Query Match             8.4%;  Score 83.6;  DB 4;  Length 506;
  Best Local Similarity  90.8%;  Pred. No. 5.8e-12;
  Matches   89;  Conservative   0;  Mismatches   9;  Indels   0;  Gaps   0;

Qy     34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     41 CGATCTCCTCACCTCCTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCA 100

Qy     94 TGAGCCACCGCGCCCGGCCCCCGACACCCAGTTTTAAA 131
          |||||||||||||||||||| ||||||||||||||||
Db    101 TGAGCCACCGCGCCCGGCCCCCAATCATCGTTTTTAAA 138


RESULT 2
US-09-918-686-2/c
; Sequence 2, Application US/09918686
; Patent No. 6475739
; GENERAL INFORMATION:
;  APPLICANT: Brunkow, Mary
;  APPLICANT: Proll, Sean
;  APPLICANT: Paeper, Bryan
;  APPLICANT: Staehling-Hampton, Karen
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING
;  TITLE OF INVENTION: GENOMIC DELETIONS
;  FILE REFERENCE: 240083.515
;  CURRENT APPLICATION NUMBER: US/09/918,686
;  CURRENT FILING DATE: 2001-07-30
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
```

; LENGTH: 51719
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1246, 2572, 2604
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-2

Query Match   8.1%; Score 80.4; DB 4; Length 51719;
Best Local Similarity 88.8%; Pred. No. 1.9e-10;
Matches   87; Conservative   0; Mismatches 11; Indels   0; Gaps   0;

```
Qy     34 CGATCTCCTGACCTCGTGATCCGCCCGCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          ||||||||||||||||||||||||| |||  |||| ||||||||||||||||| ||||||||
Db   1119 CGATCTCCTGACCTCGTGATCACCCGGCCTCGGCCTCCCAAAGTGCTGGGTTACAGGCA 1060

Qy     94 TGAGCCACCGGCCCGGCCCCCGACACCTAGTTTTAAA 131
          |||||||||| |  ||||||||
Db   1059 CGAGCCACCGCACCCGGCCCCCCATATCTATTTTTTAA 1022
```

RESULT 3
US-09-918-686-1/c
; Sequence 1, Application US/09918686
; Patent No. 6475739
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Proll, Sean
; APPLICANT: Paeper, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: GENOMIC DELETIONS
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918,686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 92139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7043, 8369, 8401
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-1

Query Match   8.1%; Score 80.4; DB 4; Length 92139;
Best Local Similarity 88.8%; Pred. No. 2.4e-10;
Matches   87; Conservative   0; Mismatches 11; Indels   0; Gaps   0;

```
Qy     34 CGATCTCCTGACCTCGTGATCCGCCCGCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          ||||||||||||||||||||||||| |||  |||| ||||||||||||||||| ||||||||
Db   6916 CGATCTCCTGACCTCGTGATCACCCGGCCTCGGCCTCCCAAAGTGCTGGGTTACAGGCA 6857

Qy     94 TGAGCCACCGGCCCGGCCCCCGACACCTAGTTTTAAA 131
          |||||||||| |  ||||||||
Db   6856 CGAGCCACCGCACCCGGCCCCCCATATCTATTTTTTAA 6819
```

RESULT 4
US-09-621-976-9083/c
; Sequence 9083, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 9083
; LENGTH: 511
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-9083

Query Match   8.0%; Score 78.8; DB 4; Length 511;
Best Local Similarity 84.0%; Pred. No. 9.4e-11;
Matches   89; Conservative   0; Mismatches 17; Indels   0; Gaps   0;

```
Qy     34 CGATCTCCTGACCTCGTGATCCGCCCGCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          ||||||||||||||| ||||| |||  |||||| ||||||||||||||||||||||||| ||
Db    439 CGATCTCCTGACCTCATGATCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 380

Qy     94 TGAGCCACCGCGCCCGGCCCCCGACACCTAGTTTAAGGGCCCCT 139
          ||||||||||||||||||||| |
Db    379 TGAGCCACCGCGCCCGGCCCCGGCCCTAAATTTGTACTTCTAAAGGGATGCT 334
```

RESULT 5
US-09-621-976-8284/c
; Sequence 8284, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8284
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8284

Query Match   7.9%; Score 78.4; DB 4; Length 480;
Best Local Similarity 88.5%; Pred. No. 1.2e-10;
Matches   85; Conservative   0; Mismatches 11; Indels   0; Gaps   0;

```
Qy     34 CGATCTCCTGACCTCGTGATCCGCCCGCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          |||||||||||||||||||||||||||  |||| ||||||||||||||||||||||||| ||
Db    472 CGATCTCCTGACCTCGTGATCCGCCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 413

Qy     94 TGAGCCACCGCGCCCGGCCCCCGACACCACTAGTTTTA 129
          ||||||||||||||||||||  ||  |||||
Db    412 TGAGCCACCGCGCCCGGCCCGGGGAGATATATTTATA 377
```

RESULT 6
US-09-269-939A-41
; Sequence 41, Application US/09269939A
; Patent No. 6635431
; GENERAL INFORMATION:
; APPLICANT: Bihain, Bernard
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Yen-Potin, Frances
; TITLE OF INVENTION: LSR Receptor, Its Activity, Its Cloning and Its Application To
; TITLE OF INVENTION: The Diagnosis Prevention And/or Treatment of Obesity and
; TITLE OF INVENTION: Related Risks or Complications
; FILE REFERENCE: GENSET.035APC
; CURRENT APPLICATION NUMBER: US/09/269,939A
; CURRENT FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: FR 97/10088
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: FR 98/05032
; PRIOR FILING DATE: 1998-04-22
; NUMBER OF SEQ ID NOS: 41

```
;   SOFTWARE: Patent.pm
;   SEQ ID NO 41
;   LENGTH: 21721
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 1898..2253
;   NAME/KEY: exon
;   LOCATION: 3438..3782
;   NAME/KEY: exon
;   LOCATION: 12064..12183
;   NAME/KEY: exon
;   LOCATION: 15049..15105
;   NAME/KEY: exon
;   LOCATION: 15670..15816
;   NAME/KEY: exon
;   LOCATION: 19486..19659
;   NAME/KEY: exon
;   LOCATION: 19806..19865
;   NAME/KEY: exon
;   LOCATION: 19963..20094
;   NAME/KEY: exon
;   LOCATION: 20236..20864
;   NAME/KEY: exon
;   LOCATION: 20954..21094
;   NAME/KEY: Misc_Feature
;   LOCATION: 715
;   OTHER INFORMATION: diverging nucleotide, G in ref genbank:AC002128
;   NAME/KEY: Misc_Feature
;   LOCATION: 1229
;   OTHER INFORMATION: diverging insertion, G in ref genbank:AC002128
;   NAME/KEY: Misc_Feature
;   LOCATION: 3676
;   OTHER INFORMATION: diverging nucleotide, T in ref genbank:AC002128
;   NAME/KEY: Misc_Feature
;   LOCATION: 5039
;   OTHER INFORMATION: diverging deletion, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 5118
;   OTHER INFORMATION: diverging nucleotide, C in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 7337
;   OTHER INFORMATION: diverging deletion, C in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 8294
;   OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 8604
;   OTHER INFORMATION: diverging nucleotide, C in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 8928
;   OTHER INFORMATION: diverging nucleotide, A in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 9021
;   OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 9851
;   OTHER INFORMATION: diverging insertion, GAATGAAA in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 9878
;   OTHER INFORMATION: diverging nucleotide, C in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 11478
;   OTHER INFORMATION: diverging nucleotide, T in ref genbank: AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 11577
;   OTHER INFORMATION: diverging deletion, C in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 11779
;   OTHER INFORMATION: diverging nucleotide, T in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 13411

;   OTHER INFORMATION: diverging deletion, T in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 13538
;   OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 13896
;   OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 14912
;   OTHER INFORMATION: diverging nucleotide, A in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 16732
;   OTHER INFORMATION: diverging nucleotide, C in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 17169
;   OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 18946
;   OTHER INFORMATION: diverging deletion, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 19474
;   OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 20500
;   OTHER INFORMATION: diverging deletion, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 20501
;   OTHER INFORMATION: diverging deletion, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 20502
;   OTHER INFORMATION: diverging deletion, A in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 21270
;   OTHER INFORMATION: diverging nucleotid, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 21356
;   OTHER INFORMATION: diverging insertion, T in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 21476
;   OTHER INFORMATION: diverging nucleotide, A in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 21588
;   OTHER INFORMATION: diverging insertion, C in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 21601
;   OTHER INFORMATION: diverging deletion, T in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 21635
;   OTHER INFORMATION: diverging insertion, G in ref genbank:AD000684
;   NAME/KEY: Misc_Feature
;   LOCATION: 19963..19965
;   OTHER INFORMATION: Potential variant splicing site AAG
;   NAME/KEY: Misc_Feature
;   LOCATION: 1..21721
;   OTHER INFORMATION: n= a,g,c or t
US-09-269-939A-41

  Query Match             7.9%;  Score 78.4;  DB 4;  Length 21721;
  Best Local Similarity   84.6%;
  Matches   88;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;
                                                        Pred. No. 4.6e-10;

Qy     34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGTTCCAAAGTGCTGGGATTACAGGCA 93
          ||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||||
Db   5178 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 5237

Qy     94 TGAGCCACCGGCGCCCGGCCCCGACACCTAGTTTTAAAGGGCCC 137
          || |||||||||||||||||  ||||||  | |||| |||
Db   5238 TGAGCCACCGGCGCCCGGCCCCTTGTTCTTAACTGTAATGCTGCC 5281

RESULT 7
US-09-269-939A-19
; Sequence 19, Application US/09269939A
```

; Patent No. 6635431
; GENERAL INFORMATION:
; APPLICANT: Bihain, Bernard
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Yen-Potin, Frances
; TITLE OF INVENTION: LsR Receptor, Its Activity, Its Cloning and Its Application To
; TITLE OF INVENTION: The Diagnosis Prevention And/or Treatment of Obesity and
; TITLE OF INVENTION: Related Risks or Complications
; FILE REFERENCE: GENSET.035APC
; CURRENT APPLICATION NUMBER: US/09/269,939A
; CURRENT FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: FR 97/10088
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: FR 98/05032
; PRIOR FILING DATE: 1998-04-22
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 22976
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1898..2253
; NAME/KEY: exon
; LOCATION: 3437..3781
; NAME/KEY: exon
; LOCATION: 12065..12184
; NAME/KEY: exon
; LOCATION: 15045..15101
; NAME/KEY: exon
; LOCATION: 15666..15812
; NAME/KEY: exon
; LOCATION: 19479..19652
; NAME/KEY: exon
; LOCATION: 19799..19858
; NAME/KEY: exon
; LOCATION: 19956..20087
; NAME/KEY: exon
; LOCATION: 20229..20854
; NAME/KEY: exon
; LOCATION: 20944..21094
; NAME/KEY: Misc_Feature
; LOCATION: 19956..19958
; OTHER INFORMATION: Potential variant splicing site AAG
US-09-269-939A-19

Query Match   7.9%;   Score 78.4;   DB 4;   Length 22976;
Best Local Similarity 84.6%;   Pred. No. 4.6e-10;
Matches 88; Conservative   0; Mismatches 16; Indels   0; Gaps   0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCCCGCCTCCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5176 CGATCTCCTGACCTCGTGATCCGCCCCGCCTCCGGCCTCCCAAAGTGCTGGGATTACAGGCG 5235

Qy   94 TGAGCCACCGCGCCCGGCCCCCGACACCTAGTTTTAAAGGGCCC 137
           |||||||||||||||||||| || ||| ||||
Db   5236 TGAGCCACCGCGCCGCCCGCCCCTTGTTCTTAACTGTAATGCTGCC 5279

RESULT 8
US-09-499-522-1
; Sequence 1, Application US/09499522
; Patent No. 6479238
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: POLYMORPHIC MARKERS OF THE LSR GENE
; FILE REFERENCE: GENSET.053AUS
; CURRENT APPLICATION NUMBER: US/09/499,522
; CURRENT FILING DATE: 2000-02-10
; EARLIER APPLICATION NUMBER: US 60/119,592

; EARLIER FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: US 60/144,784
; EARLIER FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 23187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2001..2356
; OTHER INFORMATION: exon1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3540..3884
; OTHER INFORMATION: exon2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 12163..12282
; OTHER INFORMATION: exon3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 15144..15200
; OTHER INFORMATION: exon4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 15765..15911
; OTHER INFORMATION: exon5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 19579..19752
; OTHER INFORMATION: exon6
; FEATURE:
; NAME/KEY: exon
; LOCATION: 19899..19958
; OTHER INFORMATION: exon7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 20056..20187
; OTHER INFORMATION: exon8
; FEATURE:
; NAME/KEY: exon
; LOCATION: 20329..20957
; OTHER INFORMATION: exon9
; FEATURE:
; NAME/KEY: exon
; LOCATION: 21047..21187
; OTHER INFORMATION: exon10
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 21168..21173
; OTHER INFORMATION: AATAAA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2000
; OTHER INFORMATION: potential 5'regulatory region
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 22324..23187
; OTHER INFORMATION: homology with USF2 gene in ref: embl Y07661
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 523..544
; OTHER INFORMATION: upstream amplification primer 17-2
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1047..1068
; OTHER INFORMATION: downstream amplification primer 17-2 , complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 946..963
; OTHER INFORMATION: upstream amplification primer 99-4576

```
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 1385..1402
; .   OTHER INFORMATION: downstream amplification primer 99-4576 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 1096..1115
; .   OTHER INFORMATION: upstream amplification primer 9-19
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 1616..1635
; .   OTHER INFORMATION: downstream amplification primer 9-19 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 1602..1621
; .   OTHER INFORMATION: upstream amplification primer 9-20
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 2074..2093
; .   OTHER INFORMATION: downstream amplification primer 9-20 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 2036..2053
; .   OTHER INFORMATION: upstream amplification primer 99-4557
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 2563..2580
; .   OTHER INFORMATION: downstream amplification primer 99-4557 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 2084..2102
; .   OTHER INFORMATION: upstream amplification primer 9-1
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 2483..2500
; .   OTHER INFORMATION: downstream amplification primer 9-1 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 2470..2489
; .   OTHER INFORMATION: upstream amplification primer 9-21
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 2062..2081
; .   OTHER INFORMATION: downstream amplification primer 9-21
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 3455..3474
; .   OTHER INFORMATION: upstream amplification primer 9-3
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 3882..3901
; .   OTHER INFORMATION: downstream amplification primer 9-3 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 4444..4463
; .   OTHER INFORMATION: upstream amplification primer 99-4558
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 4336..4356
; .   OTHER INFORMATION: downstream amplification primer 99-4558 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 4902..4920
; .   OTHER INFORMATION: upstream amplification primer 99-14419
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 3775..3792
; .   OTHER INFORMATION: downstream amplification primer 99-14419
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 6638..6655
; .   OTHER INFORMATION: upstream amplification primer 99-4577
;     FEATURE:

;     NAME/KEY: primer_bind
; .   LOCATION: 7072..7089
; .   OTHER INFORMATION: downstream amplification primer 99-4577 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 7995..8012
; .   OTHER INFORMATION: upstream amplification primer 99-4559
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 8576..8593
; .   OTHER INFORMATION: downstream amplification primer 99-4559 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 9622..9639
; .   OTHER INFORMATION: upstream amplification primer 99-3148
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 10023..10040
; .   OTHER INFORMATION: downstream amplification primer 99-3148 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 9964..9981
; .   OTHER INFORMATION: upstream amplification primer 99-4560
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 10546..10563
; .   OTHER INFORMATION: downstream amplification primer 99-4560 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 10996..11015
; .   OTHER INFORMATION: upstream amplification primer 99-14411 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 10492..10512
; .   OTHER INFORMATION: downstream amplification primer 99-14411 , complement
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 11972..11990
; .   OTHER INFORMATION: upstream amplification primer 99-4561
;     FEATURE:
; .   NAME/KEY: primer_bind
; .   LOCATION: 12481..12501
; .   OTHER INFORMATION: downstream amplification primer 99-4561 , complement

Query Match          7.9%;  Score 78.4;  DB 4;  Length 23187;
Best Local Similarity  84.6%;  Pred. No. 4.7e-10;
Matches   88;  Conservative   0;  Mismatches  16;  Indels   0;  Gaps   0;

Qy    34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
         |||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||
Db  5280 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 5339

Qy    94 TGAGCCACCGGCGCCCGGCCCCGACACCTAGTTTTAAAGGGCCC 137
         |||||||||||||||||||||| ||| ||
Db  5340 TGAGCCACCGGCGCCCGGCCCCTTGTTCTTAACTGTAATGCTGCC 5383

RESULT 9
US-09-621-976-12615/c
; Sequence 12615, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;   APPLICANT: Dumas Milne Edwards, J.B.
;   APPLICANT: Jobert, S.
;   APPLICANT: Giordano, J.Y.
;   TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;   FILE REFERENCE: GENSET.054PR2
;   CURRENT APPLICATION NUMBER: US/09/621,976
;   CURRENT FILING DATE: 2000-07-21
;   NUMBER OF SEQ ID NOS: 19335
;   SOFTWARE: Patent.pm
; SEQ ID NO 12615
; LENGTH: 397
```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-12615

```
Query Match             7.9%;    Score  77.8;  DB 4;  Length 397;
Best Local Similarity  87.6%;    Pred. No. 1.5e-10;
Matches   85; Conservative    0; Mismatches 12; Indels    0; Gaps    0;

Qy      34 CGATCTCCTGACCTCGTGATCCGCCGCCCCCGCCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
           ||||||||||||| ||||||||||| |||||||||| || |||||||||||||||||||||||||||
Db     194 CGATCTCCTGACCTTGTGATCCGCCGCCCCCGCCCTCCCCAAAGTGCTGGGATTACAGGCG 135

Qy      94 TGAGCCACCGCGCCCGGCCCGGCCCCCGACACCTAGTTTTAA 130
           ||||||||||||||||||||||  |||||
Db     134 TGAGCCACCGCGCCCGGCCCGGCCCCCGCCGCCCCGCATTCATTTTCTTAA 98
```

RESULT 10
US-09-702-705-321
; Sequence 321, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   APPLICANT: Fan, Liqun
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C14
;   CURRENT FILING DATE: 2000-10-30
;   NUMBER OF SEQ ID NOS: 1833
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 321
;   LENGTH: 2280
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-702-705-321

```
Query Match             7.8%;    Score  77.6;  DB 4;  Length 2280;
Best Local Similarity  82.4%;    Pred. No. 3.2e-10;
Matches   89; Conservative    0; Mismatches 19; Indels    0; Gaps    0;

Qy      34 CGATCTCCTGACCTCGTGATCCGCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
           |||||||||||||||||||||| ||| ||||||||||||||||||||||||||||||||
Db    1081 CGATCTCCTGACCTCGTGATCCGCCGCCCGGCCTCCCGACACCTAGTTTAAAGGGCCCCTGC 1140

Qy      94 TGAGCCACCGCGCCCGGCCCGGCCTGCCCCGGCCTGATTCTCTTAAAATTGAAGAGGTGCTGC 141
           ||||||||||  ||||| |||  |||||||||
Db    1141 TGAGCCACCGTGCCCGGCCGGCCTGATTCTCTTAAAATTGAAGAGGTGCTGC 1188
```

RESULT 11
US-09-736-457-321
; Sequence 321, Application US/09736457
; Patent No. 6500448
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   APPLICANT: Fan, Liqun
;   APPLICANT: Wang, Aijun
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C15
;   CURRENT APPLICATION NUMBER: US/09/736,457
;   CURRENT FILING DATE: 2000-12-13
;   SOFTWARE: FastSEQ for Windows Version 3.0
;   NUMBER OF SEQ ID NOS: 1864
; SEQ ID NO 321
;   LENGTH: 2280
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-321

```
Query Match             7.8%;    Score  77.6;  DB 4;  Length 2280;
Best Local Similarity  82.4%;    Pred. No. 3.2e-10;
Matches   89; Conservative    0; Mismatches 19; Indels    0; Gaps    0;

Qy      34 CGATCTCCTGACCTCGTGATCCGCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
           |||||||||||||||||||||| ||| ||||||||||||||||||||||||||||||||
Db    1081 CGATCTCCTGACCTCGTGATCCGCCGCCCGGCCTCCCGACACCTAGTTTTAAAGGGCCCCTGC 1140

Qy      94 TGAGCCACCGCGCCCGGCCCGGCCTGCCCCGGCCTGATTCTCTTAAAATTGAAGAGGTGCTGC 141
           ||||||||||  ||||| |||  |||||||||
Db    1141 TGAGCCACCGTGCCCGGCCGGCCTGATTCTCTTAAAATTGAAGAGGTGCTGC 1188
```

RESULT 12
US-09-614-124B-321
; Sequence 321, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C9
;   CURRENT APPLICATION NUMBER: US/09/614,124B
;   CURRENT FILING DATE: 2001-07-11
;   NUMBER OF SEQ ID NOS: 1668
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 321
;   LENGTH: 2280
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-614-124B-321

```
Query Match             7.8%;    Score  77.6;  DB 4;  Length 2280;
Best Local Similarity  82.4%;    Pred. No. 3.2e-10;
Matches   89; Conservative    0; Mismatches 19; Indels    0; Gaps    0;

Qy      34 CGATCTCCTGACCTCGTGATCCGCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
           |||||||||||||||||||||| ||| ||||||||||||||||||||||||||||||||
Db    1081 CGATCTCCTGACCTCGTGATCCGCCGCCCGGCCTCCCGACACCTAGTTTTAAAGGGCCCCTGC 1140

Qy      94 TGAGCCACCGCGCCCGGCCCGGCCTGCCCCGGCCTGATTCTCTTAAAATTGAAGAGGTGCTGC 141
           ||||||||||  ||||| |||  |||||||||
Db    1141 TGAGCCACCGTGCCCGGCCGGCCTGATTCTCTTAAAATTGAAGAGGTGCTGC 1188
```

RESULT 13
US-09-671-325-321
; Sequence 321, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary

; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 321
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-321

Query Match   7.8%; Score 77.6; DB 4; Length 2280;
Best Local Similarity 82.4%; Pred. No. 3.2e-10;
Matches   89; Conservative   0; Mismatches   19; Indels   0; Gaps   0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          ||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||||
Db   1081 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 1140

Qy   94 TGAGGCACCACCGCGCCCGGCCCCCGACACCTAGTTTTAAAGGGCCCCTGC 141
          |||||||||| |||| ||||| | |||| ||| ||||
Db   1141 TGAGGCACCGGTGCCCGGCCTGATTCTCTTAAAATTGAAGAGGTGCTGC 1188

RESULT 14
US-09-589-184-321
; Sequence 321, Application US/09589184
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 321
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-589-184-321

Query Match   7.8%; Score 77.6; DB 4; Length 2280;
Best Local Similarity 82.4%; Pred. No. 3.2e-10;
Matches   89; Conservative   0; Mismatches   19; Indels   0; Gaps   0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          ||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||||
Db   1081 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 1140

Qy   94 TGAGGCACCACCGCGCCCGGCCCCCGACACCTAGTTTTAAAGGGCCCCTGC 141
          |||||||||| |||| ||||| | |||| ||| ||||
Db   1141 TGAGGCACCGGTGCCCGGCCTGATTCTCTTAAAATTGAAGAGGTGCTGC 1188

RESULT 15
US-09-539-333D-1/c
; Sequence 1, Application US/09539333D

; Patent No. 6476208
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bihain, Bernard
; APPLICANT: Essioux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: GENSET.047AUS
; CURRENT APPLICATION NUMBER: US/09/539,333D
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: US 09/416,384
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 31..1107
; OTHER INFORMATION: 5'regulatory region g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1108..1289
; OTHER INFORMATION: exon A g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 14877..14920
; OTHER INFORMATION: exon B g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 18778..18862
; OTHER INFORMATION: exon Bbis g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25593..25740
; OTHER INFORMATION: exon C g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29388..29502
; OTHER INFORMATION: exon D g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29967..30282
; OTHER INFORMATION: exon E g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 64666..64812
; OTHER INFORMATION: exon F g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 65505..65853
; OTHER INFORMATION: exon G g35018 gene
; FEATURE:

```
  NAME/KEY: misc_feature
  LOCATION: 65854..67854
  OTHER INFORMATION: 3'regulatory region g35018 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 94124..94964
  OTHER INFORMATION: exon g35017
FEATURE:
  NAME/KEY: exon
  LOCATION: 201188..201234
  OTHER INFORMATION: exon S g35030 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 214676..214793
  OTHER INFORMATION: exon T g35030 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 215702..215746
  OTHER INFORMATION: exon U g35030 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 216836..216915
  OTHER INFORMATION: exon V g35030 gene
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 213818..215818
  OTHER INFORMATION: 3'regulatory region g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 215819..215941
  OTHER INFORMATION: exon R complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 215819..215975
  OTHER INFORMATION: exon Rbis complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 216661..216952
  OTHER INFORMATION: exon Qbis complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 216661..217061
  OTHER INFORMATION: exon Q complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 217027..217061
  OTHER INFORMATION: exon Q1 complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 229647..229742
  OTHER INFORMATION: exon X complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 230408..230721
  OTHER INFORMATION: exon P complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 231272..231412
  OTHER INFORMATION: exon Obis complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 231787..231880
  OTHER INFORMATION: exon O2 complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 231870..231879
  OTHER INFORMATION: exon O1 complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 234174..234321
  OTHER INFORMATION: exon O complement g34872 gene
FEATURE:
  NAME/KEY: exon

  LOCATION: 237406..237428
  OTHER INFORMATION: exon Nbis complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 239719..239807
  OTHER INFORMATION: exon N2 complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 239719..239853
  OTHER INFORMATION: exon N complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 240528..240569
  OTHER INFORMATION: exon M1117 complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 240528..240596
  OTHER INFORMATION: exon M1090 complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 240528..240617
  OTHER INFORMATION: exon M1069 complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 240528..240644
  OTHER INFORMATION: exon MS2 complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 240528..240824
  OTHER INFORMATION: exon M862 complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 240528..240994
  OTHER INFORMATION: exon M692 complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 240528..241685
  OTHER INFORMATION: exon M1 complement g34872 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 240800..240993
  OTHER INFORMATION: exon MS1 complement g34872 gene
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 241686..243685
  OTHER INFORMATION: 5'regulatory region g34872 gene
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 290652..292652
  OTHER INFORMATION: 3'regulatory region g34665 gene
FEATURE:
  NAME/KEY: exon
  LOCATION: 292653..292841
```

```
Query Match          7.8%;  Score 77.6;  DB 4;  Length 319608;
Best Local Similarity 86.0%; Pred. No. 1.9e-09;
Matches  86;  Conservative  0;  Mismatches  14;  Indels  0;  Gaps  0;

Qy     34 CGATCTCCTGACCTCGTGATCCGCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
          |||||||||||||||||||||||||||||| || |||||||||||||||||||||||||
Db 261651 CGATCTCCTGACCTCGTGATCCGCCGCCTCCAGCCTCCCAAAGTGCTGGGATTACAGGCG 261592

Qy     94 TGAGCCACCGCGCCCGGCCCCGACACCTAGTTTAAAGG 133
          |||| ||||| ||| |||  || |||||| | ||
Db 261591 TGACCGCCGCCGCCGGCCCGCCGCAAGTATTAATTTTTTAAGG 261552
```

Search completed: July 10, 2004, 11:58:32
Job time : 106 secs

```
                GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   July 10, 2004, 11:04:44 ; Search time 537 Seconds
                                    (without alignments)
                             8985.372 Million cell updates/sec

Title:           US-09-972-032-1
Perfect score:   990
Sequence:        1 ggaattgttctcgaggccaa........aaaatgagcggccgcaagtt 990

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        3183909 seqs, 243641669 residues

Total number of hits satisfying chosen parameters:    6357818

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
   1: /cgn2_6/ptcdata/1/pubpna/US07_PUBCOMB.seq:*
   2: /cgn2_6/ptcdata/1/pubpna/PCT_NEW_PUB.seq:*
   3: /cgn2_6/ptcdata/1/pubpna/US06_NEW_PUB.seq:*
   4: /cgn2_6/ptcdata/1/pubpna/US06_PUBCOMB.seq:*
   5: /cgn2_6/ptcdata/1/pubpna/US07_NEW_PUB.seq:*
   6: /cgn2_6/ptcdata/1/pubpna/PCTUS_PUBCOMB.seq:*
   7: /cgn2_6/ptcdata/1/pubpna/US08_NEW_PUB.seq:*
   8: /cgn2_6/ptcdata/1/pubpna/US08_PUBCOMB.seq:*
   9: /cgn2_6/ptcdata/1/pubpna/US09A_PUBCOMB.seq:*
  10: /cgn2_6/ptcdata/1/pubpna/US09B_PUBCOMB.seq:*
  11: /cgn2_6/ptcdata/1/pubpna/US09C_PUBCOMB.seq:*
  12: /cgn2_6/ptcdata/1/pubpna/US09_NEW_PUB.seq:*
  13: /cgn2_6/ptcdata/1/pubpna/US09_NEW_PUB2:*
  14: /cgn2_6/ptcdata/1/pubpna/US10A_PUBCOMB.seq:*
  15: /cgn2_6/ptcdata/1/pubpna/US10B_PUBCOMB.seq:*
  16: /cgn2_6/ptcdata/1/pubpna/US10C_PUBCOMB.seq:*
  17: /cgn2_6/ptcdata/1/pubpna/US10_NEW_PUB.seq:*
  18: /cgn2_6/ptcdata/1/pubpna/US60_NEW_PUB.seq:*
  19: /cgn2_6/ptcdata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	990	100.0	990	13	US-09-972-032-1	Sequence 1, Appli
2	86.6	8.7	631	13	US-10-027-632-251882	Sequence 251882,
3	86.6	8.7	631	16	US-10-027-632-251882	Sequence 251882,
4	85	8.6	96597	16	US-10-085-117-112	Sequence 112, App
5	84.8	8.6	99291	17	US-10-322-281-744	Sequence 744, App
6	84.2	8.5	67783	13	US-09-997-722-238	Sequence 238, App
7	83.6	8.4	2105	13	US-10-296-115-330	Sequence 330, App
8	83.6	8.4	325348	16	US-10-085-117-358	Sequence 358, App
9	83.2	8.4	636	13	US-10-027-632-269840	Sequence 269840,
10	83.2	8.4	636	16	US-10-027-632-269840	Sequence 269840,
11	83	8.4	598	13	US-10-027-632-115506	Sequence 115506,
12	83	8.4	598	16	US-10-027-632-115506	Sequence 115506,
13	83	8.4	13695	16	US-10-292-798-905	Sequence 905, App
14	83	8.4	65387	17	US-10-322-281-332	Sequence 332, App

15	82.8	8.4	790	13	US-10-027-632-125805	Sequence 125805,
16	82.8	8.4	790	13	US-10-027-632-125805	Sequence 125805,
17	82.8	8.4	790	16	US-10-027-632-125805	Sequence 125805,
18	82.8	8.4	790	16	US-10-027-632-125806	Sequence 125806,
19	82.2	8.3	741	13	US-10-027-632-149163	Sequence 149163,
20	82.2	8.3	741	16	US-10-027-632-149163	Sequence 149163, A
21	82.2	8.3	2114	13	US-10-027-632-97278	Sequence 97278, A
22	82.2	8.3	2114	16	US-10-027-632-97278	Sequence 97278, A
23	82.2	8.3	2390	9	US-10-104-047-1390	Sequence 1390, Ap
24	82.2	8.3	12477	9	US-09-764-870-559	Sequence 559, App
25	82.2	8.3	12477	9	US-09-764-853-851	Sequence 851, App
26	82.2	8.3	12477	15	US-10-125-540-559	Sequence 559, App
27	82.2	8.3	12477	15	US-10-103-313-579	Sequence 579, App
28	82.2	8.3	12477	16	US-10-158-057-339	Sequence 339, App
29	82.2	8.3	26332	13	US-10-087-192-1558	Sequence 1558, Ap
30	82	8.3	636	13	US-10-027-632-269839	Sequence 269839,
31	82	8.3	636	16	US-10-027-632-269839	Sequence 269839,
32	82	8.3	2494	13	US-10-108-260A-127	Sequence 127, App
33	82	8.3	2545	13	US-10-027-632-112152	Sequence 112152,
34	82	8.3	2545	13	US-10-027-632-112153	Sequence 112153,
35	82	8.3	2545	16	US-10-027-632-112152	Sequence 112152,
36	82	8.3	2545	16	US-10-027-632-112153	Sequence 112153,
37	82	8.3	2904	16	US-10-104-047-1356	Sequence 1356, Ap
38	82	8.3	16491	10	US-09-764-891-5505	Sequence 5505, Ap
39	81.6	8.3	168821	13	US-10-087-192-622	Sequence 622, App
40	81.6	8.2	387	9	US-09-867-701-10757	Sequence 10757, A
41	81.6	8.2	636	13	US-10-027-632-275829	Sequence 275829,
42	81.6	8.2	636	16	US-10-027-632-275829	Sequence 275829,
43	81.2	8.2	351	9	US-09-867-701-9314	Sequence 9314, Ap
44	81.2	8.2	361	9	US-09-867-701-9326	Sequence 9326, Ap
45	81.2	8.2	389	9	US-09-867-701-10783	Sequence 10783, A

ALIGNMENTS

RESULT 1
US-09-972-032-1
; Sequence 1, Application US/09972032
; Publication No. US2002008361A1
; GENERAL INFORMATION:
; APPLICANT: Case Western Reserve University
; APPLICANT: Montano, Monica
; APPLICANT: Sutton, Amelia
; TITLE OF INVENTION: A Modulator of Antiestrogen Pharmacology
; FILE REFERENCE: 27708/04003
; CURRENT APPLICATION NUMBER: US/09/972,032
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/238,190
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-972-032-1

  Query Match   100.0%; Score 990; DB 13; Length 990;
  Best Local Similarity   100.0%; Pred. No. 2.4e-269;
  Matches 990; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 GGAATTGTTCTCGAGGCCAAGAATTCGGCAAGGCCGATCTCCTGACCTCGTGATCCGCCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGAATTGTTCTCGAGGCCAAGAATTCGGCAAGGCCGATCTCCTGACCTCGTGATCCGCCCG 60

Qy   61 CCTCAGCTTCCCAAAGTGCTGGGATTACAGGCATGAGCACCACCGCCCGGCCACCACACAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCTCAGCTTCCCAAAGTGCTGGGATTACAGGCATGAGCACCACCGCCCGGCCACCACACAC 120

Qy   121 CTAGTTTTAAAGGGCCCCTGCTGTTGCTGCCGCCTCCGCCGGCCGCTCCCAGCTGCCAGTCTG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTAGTTTTAAAGGGCCCCTGCTGTTGCTGCCGCCTCCGCCGGCCGCTCCCAGCTGCCAGTCTG 180
```

```
Qy   181  GCGGGCTCAGTCCGCGGTTGCCATGTGTGGGAGGACCGCTGCGTAAGCGCTGGGATGTGG  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  GCGGGCTCAGTCCGCGGTTGCCATGTGTGGGAGGACCGCTGCGTAAGCGCTGGGATGTGG  240

Qy   241  CTTCGCTGATGCACATTGGACCGGGCTCTGACTGGGCTAGGGGAAGGGCAGGGGACGGGCGG  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  CTTCGCTGATGCACATTGGACCGGGCTCTGACTGGGCTAGGGGAAGGGCAGGGGACGGGCGG  300

Qy   301  AATTGGGCCCGAGGGGCCAGGGCCTGCCGCGACCCCGACTGCGCCTCCCGGTTGGCCCCGCAG  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  AATTGGGCCCGAGGGGCCAGGGCCTGCCGCGACCCCGACTGCGCCTCCCGGTTGGCCCCGCAG  360

Qy   361  CGCCTCCCGGTTGGCCCTGGAGTGCAGGTCTTACCGTCCGAGATCGTTCCGCAACTGGGCGA  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CGCCTCCCGGTTGGCCCTGGAGTGCAGGTCTTACCGTCCGAGATCGTTCCGCAACTGGGCGA  420

Qy   421  GCTGTGCATGGGGCGTGGCTAAGGCCGTGGTTTGGTTACGATTGGCCAGCGGACTTAAG  480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GCTGTGCATGGGGCGTGGCTAAGGCCGTGGTTTGGTTACGATTGGCCAGCGGGACTTAAG  480

Qy   481  TGTTGTCTCTGAAGAGCATGGACATTAGTCTGGAGGGTCCTGGAAGCGGGTCTCACCCGCCC  540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TGTTGTCTCTGAAGAGCATGGACATTAGTCTGGAGGGTCCTGGAAGCGGGTCTCACCCGCCC  540

Qy   541  CACCATCAAATGCGCTTAGGTCTTAGGAAGCGGGTGTGGGTGGGTGGGGCCTTAGGGGCGAGGGCG  600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  CACCATCAAATGCGCTTAGGTCTTAGGAAGCGGGTGTGGGTGGGTGGGGCCTTAGGGGCGAGGGCG  600

Qy   601  CAGACATACCCCGAAGTGGTTGGATTGTATACCGCAAGGGGCTGATCGAACCCCCCAAA  660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CAGACATACCCCGAAGTGGTTGGATTGTATACCGCAAGGGGCTGATCGAACCCCCCAAA  660

Qy   661  GACACTGGAAGGCTGTGTGTGGCTGAGGAGGGCCCGCAATCCAGTGTGTCGTGGGCTTTAC  720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  GACACTGGAAGGCTGTGTGTGGCTGAGGAGGGCCCGCAATCCAGTGTGTCGTGGGCTTTAC  720

Qy   721  AGGAAAGAGCTTCCACCTTCTTGGAGTGTGCAGATGCGATCTAGGTGTGTCCACCCGATGG  780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  AGGAAAGAGCTTCCACCTTCTTGGAGTGTGCAGATGCGATCTAGGTGTGTCCACCCGATGG  780

Qy   781  GAGCTGCGGGGCCGGGCAGATGCTGCCCCAGTACAAAGCTGATTTGGACCCTGGGGCCTCTG  840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GAGCTGCGGGGCCGGGCAGATGCTGCCCCAGTACAAAGCTGATTTGGACCCTGGGGCCTCTG  840

Qy   841  GACTTCCCTGATTCTCTGCTTCCTGCTTCCTGCCTTCAGAGTAAAATTGCAAGATCTGTCTCCGTTGGCTGCCTTCAT  900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  GACTTCCCTGATTCTCTGCTTCCTGCTTCCTGCCTTCAGAGTAAAATTGCAAGATCTGTGGTGGTCAAAAAAAAA  900

Qy   901  CCACTCTCACTTCTCTGCCTTCAGAGTAAAATTGCAAGATCTGTGGTGGTCAAAAAAAAA  960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  CCACTCTCACTTCTCTGCCTTCAGAGTAAAATTGCAAGATCTGTGGTGGTCAAAAAAAAA  960

Qy   961  AAAAAAAAAATGAGCGCGCCAAGTT  990
               |||||||||||||||||||||||||
Db   961  AAAAAAAAAATGAGCGCGCCAAGTT  990
```

RESULT 2
US-10-027-632-251882
; Sequence 251882, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, David G.
;   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;   TITLE OF INVENTION: Polymorphisms in the Human Genome
;   FILE REFERENCE: 108827.129
;   CURRENT APPLICATION NUMBER: US/10/027,632
;   CURRENT FILING DATE: 2002-04-30
;   PRIOR APPLICATION NUMBER: US 60/218,006
;   PRIOR FILING DATE: 2000-07-12
;   PRIOR APPLICATION NUMBER: US 60/198,676
;   PRIOR FILING DATE: 2000-04-20
;   PRIOR APPLICATION NUMBER: US 60/193,483
;   PRIOR FILING DATE: 2000-03-29
;   PRIOR APPLICATION NUMBER: US 60/185,218
;   PRIOR FILING DATE: 2000-02-24
;   PRIOR APPLICATION NUMBER: US 60/167,363
;   PRIOR FILING DATE: 1999-11-23
;   PRIOR APPLICATION NUMBER: US 60/156,358
;   PRIOR FILING DATE: 1999-09-28
;   PRIOR APPLICATION NUMBER: US 60/146,002
;   PRIOR FILING DATE: 1999-08-09
;   NUMBER OF SEQ ID NOS: 325720
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 251882
;   LENGTH: 631
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-251882

Query Match   8.7%; Score 86.6; DB 13; Length 631;
Best Local Similarity   91.1%; Pred. No. 3.3e-14;
  Matches   92; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy   34 CGATCTCCTGACCTCGTGATCCCGCCCCGACCTTCCCAAAGTGCTGGGATTACAGGCA 93
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   409 CGATCTCCTGACCTCGTGATCCCGCCCCGACCTTCCCAAAGTGCTGGGATTACAGGCT 468

Qy   94 TGAGCCACCGCGCCCGGCCCCCGACACACTTAGTTTTAAAGGG 134
              |||||||||||||||||||||||||||||||||||||||||||
Db   469 TGAGCCACCGCGCCCGGCCCCGCCACTGACAAGCAGTTTAAATGG 509

RESULT 3
US-10-027-632-251882
; Sequence 251882, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;   APPLICANT: Wang, David G.
;   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;   TITLE OF INVENTION: Polymorphisms in the Human Genome
;   FILE REFERENCE: 108827.129
;   CURRENT APPLICATION NUMBER: US/10/027,632
;   CURRENT FILING DATE: 2002-04-30
;   PRIOR APPLICATION NUMBER: US 60/218,006
;   PRIOR FILING DATE: 2000-07-12
;   PRIOR APPLICATION NUMBER: US 60/198,676
;   PRIOR FILING DATE: 2000-04-20
;   PRIOR APPLICATION NUMBER: US 60/193,483
;   PRIOR FILING DATE: 2000-03-29
;   PRIOR APPLICATION NUMBER: US 60/185,218
;   PRIOR FILING DATE: 2000-02-24
;   PRIOR APPLICATION NUMBER: US 60/167,363
;   PRIOR FILING DATE: 1999-11-23
;   PRIOR APPLICATION NUMBER: US 60/156,358
;   PRIOR FILING DATE: 1999-09-28
;   PRIOR APPLICATION NUMBER: US 60/146,002
;   PRIOR FILING DATE: 1999-08-09
;   NUMBER OF SEQ ID NOS: 325720
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 251882
;   LENGTH: 631
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-251882

Query Match   8.7%; Score 86.6; DB 16; Length 631;
Best Local Similarity   91.1%; Pred. No. 3.3e-14;
  Matches   92; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy   34 CGATCTCCTGACCTCGTGATCCCGCCCCGACCTTCCCAAAGTGCTGGGATTACAGGCA 93
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   409 CGATCTCCTGACCTCGTGATCCCGCCCCGACCTTCCCAAAGTGCTGGGATTACAGGCT 468

Qy   94 TGAGCCACCGCGCCCGGCCCCCGACACACTTAGTTTTAAAGGG 134

```
Db    469  TGAGGCACCGCGGGCCCCGGCCACTGACAAGCAGTTTAAATGG  509
```

RESULT 4
US-10-085-117-112
; Sequence 112, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 52945200121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 112
; LENGTH: 96597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-112

Query Match   8.6%;   Score 85; DB 16; Length 96597;
Best Local Similarity 77.4%; Pred. No. 4.5e-13;
Matches 103; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCAAAGTGCTGGGATTACAGGCA 93
            |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||||
Db   3873 CGATCTCCTGACCTCGTGATCCGCCCGCCTCCGCCCTCCCAAAGTGCTGGGATTACAGGCG 3932

Qy   94 TGAGCCACCGCGCCCGGCCCCGACACCTAGTTTTAAAGGGGCCCCCTGCTGTGTGCTGCCGC 153
            |||||||||| ||||||| ||| || ||| ||||||||| |
Db   3933 CGAGCCACCGCGGGCCCCGGCCCTTCAGCTCTACTTTGAAACATGCCCAGAATCTGCCAGATC 3992

Qy   154 TGCCGGCGCTCCC 166
            ||| ||| |||||
Db   3993 CCACACCGCTCCC 4005
```

RESULT 5
US-10-322-281-744
; Sequence 744, Application US/10322281
; Publication No. US2004012676A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 52945200100
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 744
; LENGTH: 99291
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-744

Query Match      8.6%;   Score 84.8;  DB 17;  Length 99291;
Best Local Similarity  70.6%;  Pred. No. 5.1e-13;
Matches  113;  Conservative  0;  Mismatches  47;  Indels  0;  Gaps  0;

Qy      35  GATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCAAAGTGCTGGGATTACAGGCAT  94
            |||||||||||||| |||||||||||||||||  |||||||||||| |||||||||||
Db   20756 GATCTCCTGACCTCATGATCCGCCCGCCTCAGCCTCCAAAGTGCTGACATTACAGGCGT  20815

Qy      95  GAGCCACCGCGCCCGGCCCCGACACCTAGTTTTAAAGGGGCCCCCTGCTGTTGCTGCCGCT  154
            ||||||||||| |||||
Db   20816 GAGCCACCGCGGGCCGCTGGCCCAGTAATCACTATTTTTTTAGGGTTTTTTTTGTTTTTTTGAGAT  20875
```

Qy   155 GCCGCGGCTCCCAGCTGCCCAGTTGGCGGGCTCAGTCCC 194
Db   20876 GGACTTCGCTCGTTGTTGCCATGCTGGAGTGCAGTGGCAC 20915

RESULT 6
US-09-997-722-238
; Sequence 238, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-7171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 238
; LENGTH: 67783
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (34024)..(34043)
; OTHER INFORMATION: "n" at positions 34024 through 34043 can be any base.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (39350)..(39564)
; OTHER INFORMATION: "n" at positions 39350 through 39564 can be any base.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (50784)..(50917)
; OTHER INFORMATION: "n" at positions 50784 through 50917 can be any base.
US-09-997-722-238

Query Match   8.5%;   Score 84.2; DB 12; Length 67783;
Best Local Similarity 81.0%; Pred. No. 6.8e-13;
Matches 98; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy   34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCAAAGTGCTGGGATTACAGGCA 93
            ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   51742 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCGGCCTCCCAAAGTGCTGGGATTACAGGCG 51801

Qy   94 TGAGCCACCGCGCCCGGCCCCGACACCTAGTTTTAAAGGGGCCCCCTGCTGTGTGCTGCCGC 153
            ||||||||||| ||||||
Db   51802 TGAGCCACCGCGGGCCGGCCCCGCCCGGCCCCAGTGCTTCTATTGCTTCTATTGTTGAGTGTTGAGTTCTGGCTCCAGT 51861

Qy   154 T 154
            |
Db   51862 T 51862

RESULT 7
US-10-296-115-330/c
; Sequence 330, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 330

```
; LENGTH: 2105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-296-115-330

Query Match      8.4%;   Score 83.6; DB 13;  Length 2105;
Best Local Similarity 90.8%;   Pred. No. 3.4e-13;
Matches  89;   Conservative  0;   Mismatches  9;   Indels  0;   Gaps  0;

Qy    34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
         ||||||||| ||||||||||||||||||||| || ||||||||||||||||||||||||||
Db   555 CGATCTCCTCACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 496

Qy    94 TGAGCCACCGCGCCCGGCCCGGCCCCCGACACCTAGTTTTAAA 131
         |||||||||||||||||||||||||    |  |||||||
Db   495 TGAGCCACCGCGCCCGGCCCGGCCCCCAATCATCTGTTTTTAAA 458
```

RESULT 8
US-10-085-117-358/c
; Sequence 358, Application US/10085117
; Publication No. US2003023233A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 52945200012l
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 358
; LENGTH: 325348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)..(325348)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-358

```
Query Match      8.4%;   Score 83.6; DB 16;  Length 325348;
Best Local Similarity 90.8%;   Pred. No. 1.6e-12;
Matches  89;   Conservative  0;   Mismatches  9;   Indels  0;   Gaps  0;

Qy    34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
         ||||||||| ||||||||||||||||||||| || ||||||||||||||||||||||||||
Db 156386 CGATCTCCTCACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 156327

Qy    94 TGAGCCACCGCGCCCGGCCCGGCCCCCGACACCTAGTTTTAAA 131
         |||||||||||||||||||||||||
Db 156326 TGAGCCACCGCGCCCGGCCCGGCCCCCTTTCTTTATTTTTAGA 156289
```

RESULT 9
US-10-027-632-269840/c
; Sequence 269840, Application US/10027632
; Publication No. US2002019837A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 269840
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269840

```
Query Match      8.4%;   Score 83.2; DB 13;  Length 636;
Best Local Similarity 89.8%;   Pred. No. 3.1e-13;
Matches  88;   Conservative  1;   Mismatches  9;   Indels  0;   Gaps  0;

Qy    34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
         ||||||||| ||||||||||||||||||||| || ||||||||||||||||||||||||||
Db   350 CGATCTCCTCACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 291

Qy    94 TGAGCCACCGCGCCCGGCCCGGCCCCCGACACCTAGTTTTAAA 131
         |||||||||||||||||||||||||
Db   290 TGAGCCACCGCGCCCGGCCCGGCCCYAATCATCTGTTTTTAAA 253
```

RESULT 10
US-10-027-632-269840/c
; Sequence 269840, Application US/10027632
; Publication No. US2003020407SA9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 269840
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269840

```
Query Match      8.4%;   Score 83.2; DB 16;  Length 636;
Best Local Similarity 89.8%;   Pred. No. 3.1e-13;
Matches  88;   Conservative  1;   Mismatches  9;   Indels  0;   Gaps  0;

Qy    34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
         ||||||||| ||||||||||||||||||||| || ||||||||||||||||||||||||||
Db   350 CGATCTCCTCACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 291

Qy    94 TGAGCCACCGCGCCCGGCCCGGCCCCCGACACCTAGTTTTAAA 131
         |||||||||||||||||||||||||
Db   290 TGAGCCACCGCGCCCGGCCCGGCCCYAATCATCTGTTTTTAAA 253
```

RESULT 11
US-10-027-632-115506/c
; Sequence 115506, Application US/10027632
; Publication No. US2002019837lA1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115506
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115506

Query Match 8.4%; Score 83; DB 13; Length 598;
Best Local Similarity 89.9%; Pred. No. 3.4e-13;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy   34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
         |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db  144 CGATCTCCTGACCTCGTGATCCGCCCGCCTCCGGCTTCCCAAAGTGCTGGGATTACAGGCG 85

Qy   94 TGAGCCACCGCGCCGGCGCCCCGACACCTAGTTTTAAAG 132
         ||||||||||| |||| ||||| || || | ||||||
Db   84 TGAGCCACCGCACCCGGCGCCCACAAAATTTATTTTAAAG 46
```

RESULT 12
US-10-027-632-115506/c
; Sequence 115506, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115506
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115506

Query Match 8.4%; Score 83; DB 16; Length 598;
Best Local Similarity 89.9%; Pred. No. 3.4e-13;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy   34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
         |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db  144 CGATCTCCTGACCTCGTGATCCGCCCGCCTCCGGCTTCCCAAAGTGCTGGGATTACAGGCG 85

Qy   94 TGAGCCACCGCGCCGGCGCCCCGACACCTAGTTTTAAAG 132
         ||||||||||| |||| ||||| || || | ||||||
Db   84 TGAGCCACCGCACCCGGCGCCCACAAAATTTATTTTAAAG 46
```

RESULT 13
US-10-292-798-905
; Sequence 905, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 08435/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 905
; LENGTH: 13695
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: (1)..(13695)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(264)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1113)..(1535)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3116)..(3248)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3899)..(4046)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5312)..(5443)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5789)..(5905)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6139)..(6184)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6427)..(6530)
; FEATURE:

```
;   NAME/KEY: CDS
;   LOCATION: (7286)..(7404)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (9308)..(9594)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (10681)..(10789)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (11610)..(11878)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (13365)..(13495)
US-10-292-798-905

Query Match        8.4%;    Score 83;  DB 16;  Length 13695;
Best Local Similarity  89.9%;   Pred. No. 9e-13;
Matches      89;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps      0;

Qy       34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      842 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCG 901

Qy       94 TGAGCCACCGCGCCCGGCCCCCGACACCTAGTTTTAAAG 132
            |||||||||||||||||||| ||||    ||||||
Db      902 TGAGCCACCGCGCCCGGCCCCTCAGTTTCCTCTTTTTATAG 940

RESULT 14
US-10-322-281-332
;   Sequence 332, Application US/10322281
;   Publication No. US20040126762A1
;   GENERAL INFORMATION:
;   APPLICANT: David W. Morris
;   APPLICANT: Marc S. Malandro
;   TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;   FILE REFERENCE: 52945Z001000
;   CURRENT APPLICATION NUMBER: US/10/322,281
;   CURRENT FILING DATE: 2002-12-17
;   NUMBER OF SEQ ID NOS: 866
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 332
;   LENGTH: 65387
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-322-281-332

Query Match        8.4%;    Score 83;  DB 17;  Length 65387;
Best Local Similarity  89.9%;   Pred. No. 1.5e-12;
Matches      89;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps      0;

Qy       34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41156 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCG 41215

Qy       94 TGAGCCACCGCGCCCGGCCCCCGACACCTAGTTTTAAAG 132
            |||||||||||||||||||| ||||    ||||||
Db    41216 TGAGCCACCGCGCCCGGCCCCTCAGTTTCCTCTTTTTATAG 41254

RESULT 15
US-10-027-632-125805
;   Sequence 125805, Application US/10027632
;   Publication No. US20020198371A1
;   GENERAL INFORMATION:
;   APPLICANT: Wang, David G.
;   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;   TITLE OF INVENTION: Polymorphisms in the Human Genome
;   FILE REFERENCE: 108827.129
;   CURRENT APPLICATION NUMBER: US/10/027,632
;   CURRENT FILING DATE: 2002-04-30
;   PRIOR APPLICATION NUMBER: US 60/218,006
```

```
;   PRIOR FILING DATE: 2000-07-12
;   PRIOR APPLICATION NUMBER: US 60/198,676
;   PRIOR FILING DATE: 2000-04-20
;   PRIOR APPLICATION NUMBER: US 60/193,483
;   PRIOR FILING DATE: 2000-03-29
;   PRIOR APPLICATION NUMBER: US 60/185,218
;   PRIOR FILING DATE: 2000-02-24
;   PRIOR APPLICATION NUMBER: US 60/167,363
;   PRIOR FILING DATE: 1999-11-23
;   PRIOR APPLICATION NUMBER: US 60/156,358
;   PRIOR FILING DATE: 1999-09-28
;   PRIOR APPLICATION NUMBER: US 60/146,002
;   PRIOR FILING DATE: 1999-08-09
;   NUMBER OF SEQ ID NOS: 325720
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 125805
;   LENGTH: 790
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-125805

Query Match        8.4%;    Score 82.8;  DB 13;  Length 790;
Best Local Similarity  88.8%;   Pred. No. 4.2e-13;
Matches      87;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps      0;

Qy       34 CGATCTCCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       53 CGATCTCCTCACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTGCGGGATTRCAGGCA 112

Qy       94 TGAGCCACCGCGCCCGGCCCCCGACACCTAGTTTTAAA 131
            |||||||:|||||||||||||||||  |||||
Db      113 TGAGCCWCCGCGCCGGCCGGCCCCCAATCATCTGTTTTTAAA 150

Search completed: July 10, 2004, 13:18:34
Job time : 541 secs
```

```
         GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   July 10, 2004, 09:31:54 ; Search time 3110 Seconds
                                     (without alignments)
                                     9505.967 Million cell updates/sec

Title:          US-09-972-032-1
Perfect score:  990
Sequence:       1 ggaattgttctcgaggccaa........aaatgagcggccgcaagtt 990

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       27551289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters:      55026578

Minimum DB seq length:  0
Maximum DB seq length:  2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      BST:*
            1:  em_estba:*
            2:  em_esthum:*
            3:  em_estin:*
            4:  em_estmu:*
            5:  em_estcv:*
            6:  em_estpl:*
            7:  em_estro:*
            8:  em_htc:*
            9:  gb_est1:*
           10:  gb_est2:*
           11:  gb_htc:*
           12:  gb_est3:*
           13:  gb_est4:*
           14:  gb_est5:*
           15:  em_estfun:*
           16:  em_estom:*
           17:  em_gss_hum:*
           18:  em_gss_inv:*
           19:  em_gss_pln:*
           20:  em_gss_vrt:*
           21:  em_gss_fun:*
           22:  em_gss_mam:*
           23:  em_gss_mus:*
           24:  em_gss_pro:*
           25:  em_gss_rod:*
           26:  em_gss_phg:*
           27:  em_gss_vrl:*
           28:  gb_gss1:*
           29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
c 1	904.6	91.4	1467	11	BC039500	BC039500 Homo sapi
c 2	793.8	80.2	1201	13	BX406138	BX406138 BX406138
c 3	719.2	72.6	1201	12	BM980506	BM980506 UI-CF-EN1
c 4	651.2	65.8	691	13	BU620427	BU620427 UI-H-FL1-
c 5	630.8	63.7	694	13	BU684594	BU684594 UI-CF-EN1
c 6	628.2	63.5	669	12	BM679577	BM679577 UI-E-E00-
c 7	603.2	60.9	632	13	BU617331	BU617331 UI-H-DF0-
8	602.6	60.9	633	14	CA313142	CA313142 UI-CF-FN0
9	587.8	59.4	623	14	CD742581	CD742581 UI-H-FT2-
c 10	576.2	58.2	628	13	BQ447041	BQ447041 UI-H-EU1-
c 11	449.6	45.4	480	10	BF509528	BF509528 UI-H-BI4-
12	428.4	43.3	462	12	BM687366	BM687366 UI-E-CQ1-
13	409.8	41.4	439	10	BE677246	BE677246 7d82e01.x
14	382.8	38.7	411	9	AI184808	AI184808 qd24g11.x
15	371	37.5	395	9	AA687318	AA687318 nv6le01.s
c 16	328	33.1	367	10	BF002915	BF002915 7g50h12.x
17	272.6	27.5	1084	12	BM806108	BM806108 AGENCOURT
c 18	269.6	27.2	300	12	BM672504	BM672504 UI-E-CQ1-
19	211	21.3	282	10	BF364571	BF364571 PM1-NN108
c 20	209	21.1	211	9	AI202243	AI202243 qs65d12.x
21	87.4	8.8	698	10	BF663723	BF663723 602145445
22	85.8	8.7	456	12	BI792391	BI792391 ic30a05.x
23	85.8	8.7	479	28	AQ260685	AQ260685 CITBI-E1-
24	85.8	8.7	804	14	CK000918	CK000918 AGENCOURT
25	85	8.6	295	10	BG010931	BG010931 IL5-GN023
26	83.6	8.4	703	12	BM993909	BM993909 UI-H-DF1-
27	83.6	8.4	748	14	CA427898	CA427898 UI-H-DF0-
28	83.6	8.4	795	12	BG420907	BG420907 602452096
29	83.6	8.4	917	13	BX454417	BX454417 BX454417
30	83.6	8.4	1002	13	BX432291	BX432291 BX432291
31	83.4	8.4	367	12	BG014629	BG014629 IL5-GN023
32	83.2	8.4	222	9	AV738722	AV738722 AV738722
33	83.2	8.4	831	10	BF964558	BF964558 602267855
34	82.6	8.3	186	9	AI053535	AI053535 qi72b05.x
35	82.6	8.3	190	9	AI053476	AI053476 qi78h07.x
36	82.6	8.3	190	9	AI053526	AI053526 qi72a07.x
37	82.6	8.3	190	9	AI053774	AI053774 qi73h02.x
38	82.6	8.3	197	9	AI581244	AI581244 tb71b06.x
39	82.6	8.3	208	9	AI251212	AI251212 qv38h08.x
40	82.6	8.3	213	9	AI254486	AI254486 qv35b03.x
41	82.6	8.3	215	9	AI792510	AI792510 qi72b05.y
42	c 82.6	8.3	221	9	AI345241	AI345241 tb81b11.x
43	82.6	8.3	225	9	AI792580	AI792580 qi78c01.Y
44	c 82.6	8.3	226	9	AI254649	AI254649 qv56h01.x
45	c 82.6	8.3	241	10	AW301538	AW301538 xs78d08.x

ALIGNMENTS

```
RESULT 1
BC039500
LOCUS       BC039500                1467 bp    mRNA    linear   HTC 04-MAR-2003
DEFINITION  Homo sapiens, clone IMAGE:5555626, mRNA.
ACCESSION   BC039500
VERSION     BC039500.1  GI:25058499
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1467)
AUTHORS     Strausberg,R.
TITLE       Direct Submission
JOURNAL     Submitted (01-NOV-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: cgapbs-r@mail.nih.gov
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site:      http://www-shgc.stanford.edu
```

```
        Contact: (Dickson, Mark) mcd@paxil.stanford.edu
        Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
        R. M.

        Clone distribution: MGC clone distribution information can be found
        through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
        Series: IRAK Plate: 88 Row: g Column: 1
        This clone has the following problem: retained intron.
FEATURES             Location/Qualifiers
     source          1..1467
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:555626"
                     /tissue_type="Uterus, leiomyosarcoma"
                     /clone_lib="NIH_MGC_71"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"

ORIGIN
Query Match     91.4%;  Score 904.6;  DB 11;  Length 1467;
Best Local Similarity   98.8%;  Pred. No. 3.1e-127;
Matches 932;  Conservative   0;  Mismatches   9;  Indels   2;  Gaps   2;

Qy     34  CGATCTCCTGACCTCGTGATCCGCCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA   93
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    504  CGATCTCCTGACCTCGTGATCCGCCCCGCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA   563

Qy     94  TGAGCCACCGCGCCCGGCCCCGACACCTAGTTTTAAAGGGCCCCTGCTGTTGCTGCCGC   153
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    564  TGAGCCACCGCGCCCGGCCCCGACACCTAGTTTTAAAGGGCCCCTGCTGTTGCTCTCCGC   623

Qy    154  TGCCGGCCGCTCCCAGTGCCCAGTCTGGCGGGGCTCAGTCCCGGTTGCCATGTGTGGGAG   213
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    624  TGCCGGCCGCTCCCAGTGCCCAGTCTGGCGGGGCTCAGTCCCGGTTGCCATGTGTGGGAG   683

Qy    214  ACCGGCGTCGGCTAAGGCCTGATGTGGCTTCGCTCGTCACATTGGACCCGGCGTCTGGAC   273
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    684  ACCGGCGTGCGCTGGATGTGGCTTCGCTGCTGTCTGTGCACATTGGACCCGGCGTCTGGAC   743

Qy    274  TGGGGCTAGGGGAAGGGCAGGAGGGCCGGAATTGGGCCCGAGGGCCAGGCCTCGCCACCCC   333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    744  TGGGGCTAGGGGAAGGGCAGGAGGGCCGGAATTGGGCCCGAGGGCCAGGCCTCGCCACCCC   803

Qy    334  CGACTGGGCTCCCGGTGGCCCGGCCCCGCAGCGCCCTCCCGGTGGCCGTGAGTGCAGGTCTTAC   393
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    804  CGACTGGGCTCCCGGTGGCCCGGCCCCGCAGCGCCCTCCCGGTGGCCGTGAGTGCAGGTCTTAC   863

Qy    394  CGTCCGAGATGTCCCAACTGGGCGAGCTGTGCATGGGCGTGGCTAAGGCCGTGGTTT   453
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    864  CGTCCGAGATGTCCCAACTGGGCGAGCTGTGCATGGGCGTGGCTAAGGCCGTGGTTT   923

Qy    454  GGTTACGATTGGCCAGCGGGACTTAAGTGTTGTCTCTGAAGACATGGACATTAGTTCTGG   513
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    924  GGTTACGATTGGCCAGCGGGACTTAAGTGTTGTCTCTGAAGACATGGACATTAGTTCTGG   983

Qy    514  AGGGTCCTGGAAGAGTGATCCCCGCCCCACCATCAAATGGCGCTTAGGTCTAGGAAGCGG   573
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    984  AGGGTCCTGGAAGAGTGATCCCCGCCCCACCATCAAATGGCGCTTAGGTCTAGGAAGCGG   1043

Qy    574  GTGTGGGTGGGGCTAGGGGAGGCGGAGGCGCAGACATACCCCGAAGTGGTTGGATTGTATACC   633
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1044  GTGTGGGTGGGGCTAGGGGAGGCGGAGGCGCAGACACACCCCGAAGTGGTTGGATTGTATACC   1103

Qy    634  GCAAGGGCTGATGGAACCCCAAAGACACTGGAAGGCTGTGTTGGCTGGAGGAGGGCCC   693
           ||||||||||| :::|::.||||:|||||||||||||:||:|:||||||||||||||
Db   1104  GCAAGGGCTGATGGAACCCCAAAGACACTGGAAGGCTGTGTTGGCTGGAGGAGGGCCC   1163

Qy    694  GGGCA-ATCCAGTGTCGTGGCTTTGTCGTGGCTTCACCTTCT-TGGAGTGTGCA   751
           ||||: :||||||||||||||||||||||||||||||||||:
Db   1164  GGCGAATCCAGTGTCGTGGCTTTGTCGTGGCTTCACCTTCTTCTCGGAGTGTGCA   1223

Qy    752  GATGCCGATCTAGGTGTGTCACCCGATGGGAGAGCTGCGGGGCCGGGCAGATGCTGCCCCAGT   811
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1224  GATGCCGATCTAGGTGTGTCACCCGATGGGAGAGCTGCGGGGCCGGGCAGATGCTGCCCCAGT   1283

Qy    812  ACAAAGCTGATTTGGACCTGGGGCCTTCTGACTTCCGTTGCTTCTGCTTGCATCTCCAG   871
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1284  ACAAAGCTGATTTGGACCTGGGGCCTTCTGACTTCCGTTGCTTCTGCTTGCATCTCCAG   1343

Qy    872  CAAGTCCTGCCGTTGGCTGCCTGCCTCATCCACTCTCTCACTTCTCTGCCTTCAGAGTAA   931
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1344  CAAGTCCTGCCGTTGGCTGCCTGCCTCATCCACTCTCTCACTTCTCTGCCTTCAGAGTAA   1403

Qy    932  AATTGCAAGATCTGTGGTGCAAAAAAAAAAAAAAA   974
           |||||||||||||||||||||
Db   1404  AATTGCAAGATCTGAAAAAAAAAAAAAAAAAAAAAAAAA   1446


RESULT 2
BX406138/c
LOCUS       BX406138              1201 bp    mRNA    linear   EST 13-MAY-2003
DEFINITION  BX406138 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
            CS0DM010YD19 3-PRIME, mRNA sequence.
ACCESSION   BX406138
VERSION     BX406138.1  GI:30648317
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 8395.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0AM010CB10NP1&cluster=8395.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0AM010CB10NP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DM010YD19"
                     /tissue_type="FETAL LIVER"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL LIVER"
                     /note="Organ: liver; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."

ORIGIN
Query Match     80.2%;  Score 793.8;  DB 13;  Length 1201;
Best Local Similarity   94.9%;  Pred. No. 1.7e-110;
Matches 838;  Conservative  29;  Mismatches  11;  Indels   5;  Gaps   5;

Qy     34  CGATCTCCTGACCTCGTGATCCGCCCCGCCTCAGCTTCCCAAAGTGCTGGATTACAGGCA   93
           ||||:|:::|::|:||||||||||||||||||||||||||||||||||||| ||||||||
Db    882  CGAKCKCCKGACCTCGTGATCCCCCGCCKTCCCAAAGKGCKGGGAKNACAGGCA   823

Qy     94  TGAGCCACCGCGCCCGGCCCCGACACCTAGTTTTAAAGGGCCCCTGCTGTTGCTGCCGC   153
           |||||||||||||||||||||||||||||||||||||||||||||||||| :||||||
Db    822  TGAGCCACCGCGCCCGGCCCCGACACCTAGTKTAAAGGGCCCCTTGCTGTTGTCTGCCSS   763

Qy    154  TGCCGGCCGCTCCCAGTGCCCAGTCTGGCGGGGCTCAGTCCCGGTTGCCATGTGTGGGAG   213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES          Location/Qualifiers
     source       1..760
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="UI-CF-EN1-add-a-08-0-UI"
                  /tissue_type="Primary Lung Cystic Fibrosis Epithelial
                  Cells"
                  /dev_stage="Adult"
                  /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                  /clone_lib="UI-CF-EN1"
                  /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                  modified polylinker; Site 1: EcoR I; Site 2: Not I;
                  UI-CF-EN1 is a normalized cDNA library containing the
                  following tissue(s): Primary Lung Cystic Fibrosis
                  Epithelial Cells. The library was constructed according to
                  Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                  1996. First strand cDNA synthesis was primed with an
                  oligo-dT primer containing a Not I site.  Double stranded
                  cDNA was ligated to an EcoR I adaptor, digested with Not
                  I, and cloned directionally into pT7T3-Pac vector. The
                  oligonucleotide used to prime the synthesis of
                  first-strand cDNA contains a library tag sequence that is
                  located between the Not I site and the (dT)18 tail. The
                  sequence tag for this library is CTGCTCAGGT.
                  TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                  6hr to LPS 24h
                  TAG_LIB=UI-CF-EN1
                  TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match        72.6%;  Score 719.2;  DB 12;  Length 760;
Best Local Similarity   99.1%;  Pred. No. 3.7e-99;
Matches   754;  Conservative   0;  Mismatches   4;  Indels   3;  Gaps   3;

Qy    213 GACCGCGCTCGCGTAAGCGCTGGATGTGGCTTCGCTGATGCACATTGGACCGGGCTCTGGA 272
          ||||||||||||||||||||||||||||| |:|||||||||||||:|||||||||||||||
Db    760 GACCGCGCTCGCGTAAGCGCTGGATGT-GCTTCGCTGATGCACATTGGACCGGGCTCTGGA 702

Qy    273 CTGGGCTAGGGGAAGGGCAGGAGGCGGAATTGGGCCCGAGGGCCTCGCCGACCC 332
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    701 CTGGGCTAGGGGAAGGGCAGGAGGCGGAATTGGGCCCGAGGGCCTCGCCGACCC 642

Qy    333 CCGACTGCGCCTCCCGGTGGCCCCGCAGCGCCCTCCCGGTGGCCCTGAGTGCAGGTCTTA 392
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    641 CCGANTGCGCCTCCCGGTGGCCCCGCAGCGCCCTCCCGGTGGCCCTGAGTGCAGGTCTTA 582

Qy    393 CCGTTCCGAGATCGTCCGCAACTGGGCGGACTCTGTGCGAGTGTGCGAGTGTGGCCGTGGTT 452
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    581 CCGTTCCGAGATCGTCCGCAACTGGGCGGACTCTGTGCGAGTGTGCGAGTGCCGTGGTT 522

Qy    453 TGGTTACGATTGGCCAGCGGACTTAAGTGTGTTCTGAAGAGACATTGGACATTAGTCTG 512
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    521 TGGTTACGATTGGCCAGCGGACTTAAGTGTTGTTCTGAAGAGACATGGACATTAGTCTG 462

Qy    513 GAGGGTCCTGGAAGAGTGATCCCCGCCCCACCATCAAATGCGCTTAGGTCTAGGAAGCG 572
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461 GAGGGTCCTGGAAGAGTGATCCCCGCCCCACCATCAAATGCGCTTAGGTCTAGGAAGCG 402

Db    762 TGCCGCCGCTCGCAGTGCCCAGTGTCTGGCGGGCTCAGTCCCGCGGTGCCATGTGTGGGAG 703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    214 ACCGCGCTCGCGTAAGCGCTGGATGTGGCTTCGCTGATGCACATTGGACCGGGCTCTGGAC 273

Db    702 ACCGCGCTCGCGTAAGCGCTGGATGTTGSTTCGCTGATGCACATTGGACCGGGCTCTGGAC 643
          ||||||||||||||||||||||||||| |:|||||||||||||||||||||||||||||||
Qy    274 TGGGCTAGGGGAAGGGCCGGAATTGGGCCGAGGGCCAGGCCTCGCCGACCCC 333

Db    642 TGGGCTAGGGGAAGGGCCAGGAGGGCGGAATT-GGCCCGAGGGCCAGGCCTCGCCGACCCC 584
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Qy    334 CGACTGGCGCCTCCCGGTGGCCCCGCAGCGCCCTCCCGGTGGCCCTGAGTGCAGGTCTTAC 393

Db    583 CGACTGCGCCTCCCGGTGGCCCCGCAGCGCCCTCCCGGTGGCCCTGAGTGCAGGTCTTAC 524
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    394 CGTTCCGAGATCGTCCGCAACTGGGCGGAGCTGTGCATGGGCGGTGGCTAAGGCCGTGGTTT 453

Db    523 CGTTCCGAGATCGTCCGCAACTGGGCGGAGCTGTGCATGGGCGGTGGCTAAGGCCGTGGTTT 464
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    454 GGTTACGATTGGCCAGCGGGGACTTAAGTGTGTTTCTGAAGAGACATGGACATTAGTCTGG 513

Db    463 GGTTACGATTGGCCAGCGGGGACTTAAGTGTGTTCTGAAGAGACATGGACATTAGTCTGG 404
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    514 AGGGTCCTGGAAGAGTGATCCCCGCCCCACCATCAAATGGCGCTTAGGTCTAGGAAGCGG 573

Db    403 AGGGTCCTGGAAGAGTGATCCCCGCCCCACCATCAAATGGCGCTTAGGTCTAGGAAGCGG 344
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    574 GTGTGGGTGGGGCCTTAGGGCGAGGCGCAGGACATACCCGAAGTGGTTGGATTGTPATACC 633

Db    343 GTGTGGGTGGGGCCTTAGGGCCGAGGCCAGGACACACACCCCGAAGTGGTTGGATTGTATACC 284
          ||||||||||||||||| ||| ||||||||||||   || |||||||||||||||||||
Qy    634 GCAAGGGGCTGGATCGGAACCCCCCAAAGACACTGGAAGGGCTGTGTGGCTGAGGAGGGCCC 693

Db    283 GCAAGGGGCTGGATCGGAACCCCCCAAAGACACTGGAAGGGCTGTGTGGCTGAGGAGGGCCC 224
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    694 GGCA-ATCCAGTGTGTCGTGGCTTTACAGGAAAAGAGCTCCACCTTCT-TGGAGTGTGCA 751

Db    223 GGCAGATCCAGTGTGTCCTGGGCTTTACAGGAAAGAGCTCCACCTTSTCTGGAGTGTGCA 164
          |||| |||||||||||| ||| ||||||||||| ||||||||||||:| ||||||||||||
Qy    752 GATGCCGATCTAGGTGTGTGTCCACCCGATGGGAGCTG-CGGGCCGGGCGCAGATGCTGCCCCAG 810

Db    163 GATGRGATCTAGGTGTGTGTCCACCCGATGGGAGCTGACGGGCCGGGCMARAWGCTGCCCCAG 104
          ||||| |||||||||||||||||||||||||| ||:|||| ||| ||:|||||||||||||
Qy    811 TACAAAGCTGATTTGGACCTGGGCCCTCTGGGACTTCCCTGACTTCCTGCTTGCATCTCCA 870

Db    103 TACAAAGCHGATTTGGACCTGGGRMMTCTGVACTTMCCTGACTCTCWGCTTGCATCTCCA 44
          ||||||||:||||||||||||||::|: :||:|||:|||||||||  :||||||||||||
Qy    871 GCAAAGT-CCTGTCCCGTTGGCTGCCTTCATCCACTCTCTCAC 912
          ::|| || ::| :|||||||:|:||::||||||:|||||||||
Db    43  GMAAAGDACMTGTCCMGWTGGAAGCCATCATCCAMTCTNTCAC 1
```

```
RESULT 3
BM980506/c
LOCUS        BM980506         760 bp    mRNA      linear   EST 21-FEB-2003
DEFINITION   UI-CF-EN1-add-a-08-0-UI s1 UI-CF-EN1 Homo sapiens cDNA clone
             UI-CF-EN1-add-a-08-0-UI 3', mRNA sequence.
ACCESSION    BM980506
VERSION      BM980506.1  GI:19602038
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 760)
  AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE      Normalization and subtraction: two approaches to facilitate gene
             discovery
  JOURNAL    Genome Res. 6 (9), 791-806 (1996)
  MEDLINE    97044477
  PUBMED     8889548
COMMENT      Contact: McCray, PB
             McCray Lab
```

```
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dt)18 tail. The sequence tag for this library is
GAGGTCGGTG. The cell lines were provided by Dr. James
Martin from the University of Iowa.
TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FL1
TAG_SEQ=GAGGTCGGTG"
```

ORIGIN

```
Query Match        65.8%;   Score 651.2;  DB 13;  Length 691;
Best Local Similarity 99.1%;   Pred. No. 6.9e-89;
Matches  686;  Conservative  0;  Mismatches  3;  Indels  3;  Gaps  3;

Qy    279 TAGGGGAAGGGACAGGAGGGCGAATTGGCCCGAGGGCCAGGCCTCGCCGACCCCCGACT 338
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    691 TAGGGGAAGGGCACGAGGGCGGAA-TGGGCCCGAGGGCCAGGCCTCGCCGACCCCCGACT 633

Qy    339 GCGCCTCCCGGTGGCCCGCAGCCGCCCGGTGGCCCTGGAGTGCAGGTCTTACCCGTCC 398
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    632 GCGCCTCCCGGTGGCCCGCAGCCGCCCGGTGGCCCTGGAGTGCAGGTCTTACCCGTCC 573

Qy    399 GAGATCGTCCGCAACTGGGCGAGCTGTGCATGGGGCGTGGCTAAGGCCGTGGTTGGTTA 458
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    572 GAGATCGTCCGCAACTGGGCGAGCTGTGCATGGGGCGTGGCTAAGGCCGTGGTTGGTTA 513

Qy    459 CGATTGGCCAGCGGGACTTAAGTGTTGTCTCTGAAGAGCATGACATTAGTCTGGAGGGT 518
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    512 CGATTGGCCAGCGGGACTTAAGTGTTGTCTCTGAAGAGCATGACATTAGTCTGGAGGGT 453

Qy    519 CCTCGAAGAGTGATCCCCGCCCCATCAAATGGCGGCTTAGGTCTAGGAAGCGGGGTGTG 578
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    452 CCTCGAAGAGTGATCCCCGCCCCACCATCAAATGGCGGCTTAGGTCTAGGAAGCGGGGTGTG 393

Qy    579 GGTGGGGGCCTTAGGGCGCAGGGCGCAGACATACCCCGAAGTGGTTGGATTGTATACCGCAAG 638
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    392 GGTGGGGGCCTTAGGGCGCAGGGCGCAGACACTGGAAGCTGTGTGGCTGAGGAGGGCCGGCA- 333

Qy    639 GGGCTGGATCGAACCCCCAAAGACACTGGAAGCTGTGTGGCTGAGGAGGGCCGGCA- 697
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    332 GGGCTGGATCGAACCCCCAAAGACACTGGAAGCTGTGTGGCTGAGGAGGGCCGGCGCAG 273

Qy    698 ATCCAGTGTGTCGTGGGCTTACAGGAAAGAGCTTCCACCTTCT-TGGAGTGTGCAGATGC 756
          ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db    272 ATCCAGTGTGTCGTGGGCTTACAGGAAAGAGCTTCCACCTTCTCTGGAGTGTGCAGATGC 213

Qy    757 GATCTAGTGTGTCCACCGATGGGAGCTGCGCGGCCGGCAGATGCTGCCCCAGTACAAA 816
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    212 GATCTAGGTGTGTCACCCGATGGGAGCTGCGCGGCCGGCAGATGCTGCCCCAGTACAAA 153

Qy    817 GCTGATTTGGACCTGGGCCTTGGACTTCCTGATTCTTGCTTGCTTTGCATCTCCAGCAAAG 876
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    152 GCTGATTTGGACCTGGGCCTTGGACTTCCTGATTCTTGCTTGCTTTGCATCTCCAGCAAAG 93

Qy    877 TCCTGTCCCGTTGGCTGCCTTCATCCACTTCTCCACTTCTCCTTCTCTGCCTTCAGAGTAAAATTG 936
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    92 TCCTGTCCCGTTGGCTGCCTTCATCCACTTCTCCACTTCTCCTTCTCTGCCTTCAGAGTAAAATTG 33

Qy    937 CAAGATCTGTGGTGCAAAAAAAAAAAAAAAAAAA 968
          ||||||||||||||||||||||
Db    32 CAAGATCTGTGGTGCAAAAAAAAAAAAAAAAAAAA 1
```

RESULT 5
BU684594/c
LOCUS   BU684594   694 bp   mRNA   linear   EST 07-OCT-2002
DEFINITION UI-CF-EN1-act-a-22-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
            UI-CF-EN1-act-a-22-0-UI 3', mRNA sequence.
ACCESSION   BU684594

```
Qy    573 GGTGTGGGTGGGGCCTTAGGGCGGAGGGCGCAGACATACCCCGAAGTGGTTGGATTGTATAC 632
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 GGTGTGGGTGGGGCCTTAGGGCGGAGGGCGCAGACACACCCCGAAGTGGTTGGATTGTATAC 342

Qy    633 CGCAAGGGGCTGATCGAACCCCCAAAGACACTGGAAGGCTGTGTGGCTGAGGAGGGCC 692
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 CGCAAGGGGCTGATCGAACCCCCAAAGACACTGGAAGGCTGTGTGGCTGAGGAGGGCC 282

Qy    693 CGGCA-ATCCAGTGTGTCGTGGGCTTTACAGGAAAGAGCTTCCACCTTCT-TGGAGTGTGC 750
          ||||| |||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    281 CGGCAGATCCAGTGTGTCGTGGGCTTTACAGGAAAGAGCTCCACCTTCTCTGGAGTGTGC 222

Qy    751 AGATGGGATCTAGGTGTGTCCACCCGATGGGAGCTGCGCGGCCGGCAGATGCTGCCCCAG 810
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 AGATGGGATCTAGGTGTGTCCACCCGATGGGAGCTGCGCGGCCGGCAGATGCTGCCCCAG 162

Qy    811 TACAAAGCTGATTTGGACCTGGGGCCTTCCCTGGACTTCCCTGATTCTTGCTTGCATCTCCA 870
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 TACAAAGCTGATTTGGACCTGGGGCCTTCCCTGGACTTCCCTGATTCTTGCTTGCATCTCCA 102

Qy    871 GCAAAGTCCTGTCCCGTTGGCTGCCTTCATCCACTTCTCTGCCTTCAGAGTA 930
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 GCAAAGTCCTGTCCCGTTGGCTGCCTTCATCCACTTCTCTGCCTTCAGAGTA 42

Qy    931 AAATTGCAAGATCTGTGGTGCAAAAAAAAAAAAAAAA 971
          |||||||||||||||||||||||||||||||||||||
Db    41 AAATTGCAAGATCTGTGGTGCTTAAAAAAAAAAAAAAAAA 1
```

RESULT 4
BU620427/c
LOCUS   BU620427   691 bp   mRNA   linear   EST 23-SEP-2002
DEFINITION UI-H-FL1-bfw-n-19-0-UI.s1 NCI_CGAP_FL1 Homo sapiens cDNA clone
            UI-H-FL1-bfw-n-19-0-UI 3', mRNA sequence.
ACCESSION   BU620427
VERSION   BU620427.1   GI:23286642
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 691)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES   Location/Qualifiers
     source   1..691
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FL1-bfw-n-19-0-UI"
                     /tissue_type="Cell lines"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_FL1"
                     /note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
                     (Pharmacia) with a modified polylinker; Site_1: EcoR I;
                     Site_2: Not I; NCI_CGAP_FL1 is a normalized cDNA library
                     derived from a pool of mRNA obtained from 4 cell lines
                     from grade III chondrosarcoma tissues. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
```

```
VERSION     BU684594.1  GI:23537704
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 694)
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
PUBMED      8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES    Location/Qualifiers
     source  1..694
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="UI-CF-EN1-act-a-22-0-UI"
             /tissue_type="Primary Lung Cystic Fibrosis Epithelial
             Cells"
             /dev_stage="Adult"
             /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
             /clone_lib="UI-CF-EN1"
             /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
             modified polylinker; Site 1: EcoR I; Site 2: Not I;
             UI-CF-EN1 is a normalized cDNA library containing the
             following tissue(s): Primary Lung Cystic Fibrosis
             Epithelial Cells. The library was constructed according to
             Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
             1996. First strand cDNA synthesis was primed with an
             oligo-dT primer containing a Not I site. Double stranded
             cDNA was ligated to an EcoR I adaptor, digested with Not
             I, and cloned directionally into pT7T3-Pac vector. The
             first-strand cDNA contains a library tag sequence that is
             located between the Not I site and the (dT)18 tail. The
             sequence tag for this library is CTGCTCAGGT.
             TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS
             6hr to LPS 24h
             TAG_LIB=UI-CF-EN1
             TAG_SEQ=CTGCTCAGGT"
ORIGIN

Query Match         63.7%;     Score 630.8; DB 13; Length 694;
Best Local Similarity 97.7%;   Pred. No. 8.3e-86;
Matches  671;   Conservative  0;  Mismatches  13;  Indels  3;  Gaps  3;

Qy  290 CAGGAGGGCGGAATTGGGCCCGAGGGCCAGGCCTGCGCACCCCCGACTGCGCCTCCCCGG 349
        |||||||||||| ||||||||||||||||||||||||||||||| ||||||||||||||
Db  694 CAGGAGGGCGGAA-TGGGCCCGAGGGCCAGGCCTGCGCACCCCCANTGCGCCTCCCCGG 636

Qy  350 TGGCCCGCAGCGCCTCCCGGTGGCCCTGGACTGGACTGGCAGGTGCAGGTCTTACCGTCCG 409
        |||||||||||||||||||||||||||||| |||  |||||||||||||||||||||
Db  635 TGGCCCGCAGCGCCTCCCGGTGGCCCTGGAGTGCGCAGGTCTTACCGTCCGAGATCGTCCG 576

Qy  410 CAACTGGCGGAGCTGTCATGGGCCGTGGCTAAGGCCGTGGTTGGTTACGATTGGCCAG 469
```

```
Db  575 CAACTGGCGGAGCTGTCATGGGGCCGGTGGCTGGCTAAGGCCGTGGTTGGTTACGATTGGCCAG 516
        ||||||||||||||||||||| ||  ||||| |||||||||||||||||||||||||||||||
Qy  470 CGGGACTTAAGTGTTGTCTCTGAAGACATTGGACATTAGTCTGGAGGGTCCTGAAGAGT 529
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  515 CGGGACTTAAGTGTTGTCTCTGAAGACATTGGACATTAGTCTGGAGGGTCCTGAAGAGT 456
Qy  530 GATCCCGCCCACCATCAAATGGCCTTAGGTCAGGAAGCGGGTGTGGGTGGGGCCTT 589
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  455 GATCCCGCCCACCATCAAATGGCCTTAGGTCAGGAAGCGGGTGTGGGTGGGGCCTT 396
Qy  590 AGGGGCCAGGGCCAGACATACCCCGAAGTGGTTGGATTGTATACCCGAAGGGGCTGATCG 649
        ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db  395 AGGGGCCAGGGCCAGACACACCCCGAAGTGGTTGGATTGTATACCCGAAGGGGCTGATCG 336
Qy  650 AACCCCCAAAGACACTGGAAGGCTGTGTGGCTGAGGAGGGCCCGGCA-ATCCAGTGTGT 708
        |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  335 AACCCCCAAAGACACTGGAAGGCTGTGTGGCTGAGGAGGGCCCGGCATCCAGTGTGT 276
Qy  709 CGTGGGCTTTACAGGAAAAGAGCTCCACCTTCT-TGGAGTGTGCGATGGCGATCTAGGTGT 767
        ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  275 CGTGGGCTTTACAGGAAAAGAGCTCCACCTTCTCTCGGAGTGTGCGATGGCGATCTAGGTGT 216
Qy  768 GTCCACCGATGGGAGCTGCGGGCGCCGGGCAGATGCTGCCCCGAGTACAAAGCTGATTTGGA 827
        ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  215 GTCCACCCGATGGGAGCTGCGGGCGCCGGGCAGATGCTGCCCCGAGTACAAAGCTGATTTGGA 156
Qy  828 CCTGGGGCCTCTGGACTTCCCTGATTCTTGCTTGCATCTCCAGCAAAGTCCTGTCCCGT 887
        ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  155 CCTGGGGCCCTCTGGACTTCCCTGATTCTTGCTTGCATCTCCAGCAAAGTCCTGTCCCGT 96
Qy  888 TGGCTGCCTTCATCCACTCCTCTCCGCCTTCAGAGTAAAATTGCAAGATCTGTG 947
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   95 TGGCTGCCTTCATCCACTCCTCTCCGCCTTCAGAGTAAAATTGCAAGATCTGTG 36
Qy  948 GTGCAAAAAAAAAAAAAAAAAAAAAAA 974
        ||||  |  ||||||||||
Db   35 GTGCTTACTGGATCTGAAAAAAAAAAA 9
```

```
RESULT 6
BM679577/c
LOCUS       BM679577               669 bp    mRNA    linear   EST 27-FEB-2002
DEFINITION  UI-E-EOO-aia-1-05-0-UI.s1 UI-E-EOO Homo sapiens cDNA clone
            UI-E-EOO-aia-1-05-0-UI 3', mRNA sequence.
ACCESSION   BM679577
VERSION     BM679577.1  GI:18989473
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 669)
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
PUBMED      8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
            cDNA library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Forward
```

FEATURES
    source   Location/Qualifiers
                    1..669
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="UI-E-E00-aia-l-05-0-UI"
                    /tissue_type="fetal eye"
                    /dev_stage="fetal"
                    /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                    /clone_lib="UI-E-E00"
                    /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                    modified polylinker; Site_1: EcoR I; Site_2: Not I;
                    UI-E-E00 is a cDNA library containing the following
                    tissue(s): fetal eye. The library was constructed
                    according to Bonaldo, Lennon and Soares, Genome Research,
                    6:791-806, 1996. First strand cDNA synthesis was primed
                    with an oligo-dT primer containing a Not I site. Double
                    stranded cDNA was ligated to an EcoR I adaptor, digested
                    with Not I, and cloned directionally into pT7T3-Pac
                    vector. The oligonucleotide used to prime the synthesis of
                    first-strand cDNA contains a library tag sequence that is
                    located between the Not I site and the (dT)18 tail. The
                    sequence tag for this library is CCGGTATACC. This library
                    was created for the program, Gene Discovery in the Visual
                    System, supported by National Eye Institute (NEI)."
                    TAG_LIB=UI-E-E00
                    TAG_SEQ=CCGGTATACC"

ORIGIN

Query Match   63.5%;   Score 628.2; DB 12; Length 669;
Best Local Similarity 98.3%;   Pred. No. 2.1e-85;
Matches 655; Conservative   0; Mismatches   9; Indels   2; Gaps   2;

Qy   311 GAGGGCCAGGCCTCGCCGCACCCCGACTGCGCTCCCGGTGCGCCCGCAGCGCCTCCCGG 370
          ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   669 GAGGGCCAGGCCTCGCCGCACCCCGANTGCGCTCCCGGTGCGCCCGCAGCGCCTCCCGG 610

Qy   371 TGGCCCTTGGAGTGCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGCGAGCTGTGCATG 430
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   609 TGGCCCTTGGAGTGCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGCGAGCTGTGCATG 550

Qy   431 GGGCGGTGGCTAAGGCCGTGGTTGGTTACGATTGGCCAGCGGGACTTAAGTGTTGTCTCT 490
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   549 GGGCGGTGGCTAAGGCCGTGGTTGGTTACGATTGGCCAGCGGGACTTAAGTGTTGTCTCT 490

Qy   491 GAAGAGCATGACATTAGTCTGGAGGGTCCTGGAAGAGTGATCCCGCCCACCATCAAA 550
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   489 GAAGAGCATGACATTAGTCTGGAGGGTCCTGGAAGAGTGATCCCGCCCACCATCAAA 430

Qy   551 TGGCGCTTAGGTCTAGGAAGCCGGTCGTGGTGGGCCTTAGGGCGGAGGGCGCAGACATACC 610
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   429 TGGCGCTTAGGTCTAGGAAGCCGGTCGTGGTGGGCCTTAGGGCGGAGGGCGCAGACACACC 370

Qy   611 CCGAAGTGGTTGGATTGTATACCGCAAGGGGCTGGATCGAACCCCCAAAGACACTGGAA 670
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   369 CCGAAGTGGTTGGATTGTATACCGCAAGGGGCTGGATCGAACCCCCAAAGACACTGGAA 310

Qy   671 GGCTGTGTGGGCTGAGGAGGGCCCGGCA-ATCCAGTGTGTCGTGGGCTTTACAGGAAAGAG 729
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   309 GGCTGTGTGGGCTGAGGAGGGCCCGGCAGATCCAGTGTGTCGTGGGCTTTACAGGAAAGAG 250

Qy   730 CTCCACCTTCT-TGGAGTGTGCAGATGCGATCTAGGTGTGTCCACCCGATGGGGAGCTGCG 788
          ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   249 CTCCACCTTCTCTGGAGTGTGCAGATGCGATCTAGGTGTGTCCACCCGATGGGGAGCTGCG 190

Qy   789 GGCCGGGCAGATGCTGCCCCAGTACACAAAGCTGATTTGGACCTGGGGCCTTCCT 848
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   189 GGCCGGGCAGATGCTGCCCCAGTACACAAAGCTGATTTGGACCTGGGGCCTTCCC 130

Qy   849 TGATTCTCTGCTTGCATCTCCAGCAAAGTCCTGCCCGTCCGGTTGGCTGCCTTCATCCACTCTC 908
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   129 TGATTCTCTGCTTGCATCTCCAGCAAAGTCCTGCCCGTCCCGTGTCCCGTGTGCCTGCCTTCATCCACTCTC 70
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||| ||||||||||||||||
Qy   909 TCACTTCTCTGCTTCAGAGTAAAATTGCAAGATCCTGGTGTGCAAAAAAAAAAAAAAA 968
          ||||||||||||||||||||||||||||||||||||||||
Db   69 TCACTTCTCTGCTTCAGAGTAAAATTGCAAGATCTGGTGGTGTGTTACTGGGAAAAAAAA 10

Qy   969 AAAAAA 974
          ||||||
Db   9 AAAAAA 4

RESULT 7
BU617331/c
LOCUS   BU617331   632 bp   mRNA   linear   EST 23-SEP-2002
DEFINITION UI-H-DF0-bep-n-09-0-UI.s1 NCI_CGAP_DF0 Homo sapiens cDNA clone
            UI-H-DF0-bep-n-09-0-UI 3', mRNA sequence.
ACCESSION   BU617331
VERSION   BU617331.1 GI:23283539
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 632)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.

FEATURES   Location/Qualifiers
    source   1..632
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-DF0-bep-n-09-0-UI"
                     /tissue_type="Subchondral Bone"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_DF0"
                     /note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_DF0 is a cDNA library containing the following
                     tissue(s): Subchondral Bone. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is GTTAAGCGTC.
                     TAG_TISSUE=subchondral bone
                     TAG_LIB=UI-H-DF0
                     TAG_SEQ=GTTAAGCGTC"

ORIGIN

Query Match   60.9%;   Score 603.2; DB 13; Length 632;
Best Local Similarity 99.2%;   Pred. No. 1.2e-81;
Matches 627; Conservative   0; Mismatches   3; Indels   2; Gaps   2;

Qy   339 GCGGCTCCCCGGTGCCCCGCCAGCGCCTCCCGGTGCCCTGGAGTGCGCAGGTGCGGCCCTGGAGTCTTACCGTCC 398
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   632 GCGGCTCCCCGGTGCCCCGCCAGCGCCTCCCGGTGCCCTGGAGTGCGCAGGTGCAGGTGCGGCCCTGGAGTCTTACCGTCC 573

```
                    Genetics (www.resgen.com) or from Open Biosystems
                    (www.openbiosystems.com).
                    Seq primer: M13 FORWARD
                    POLYA=Yes.

FEATURES             Location/Qualifiers
     source          1..633
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-FN0-aex-n-23-0-UI"
                     /tissue_type="Human Lung Epithelial cells"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-CF-FN0"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site 1: EcoR I; Site 2: Not I;
                     UI-CF-FN0 is a subtracted cDNA library derived from two
                     normalized Human lung epithelial cell libraries (EN1 and
                     DU1) The library was subtracted according to according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. For additional information, contact:
                     bento-soares@uiowa.edu
                     TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                     6hr to LPS 24h
                     TAG_LIB=UI-CF-FN0
                     TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match          60.9%;   Score 602.6;  DB 14;  Length 633;
Best Local Similarity 99.1%;  Pred. No. 1.5e-81;
Matches 627;  Conservative 0;  Mismatches 0;  Indels 4;  Gaps 2;
```

```
Qy 341 GCCTCCCGGTGGCCCCGCAGCGCCTCCCGGTGGCCCCGCAGCGCCTCCCGGTGGCCGTCTTACCGTCCGA 400
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 633 GCCTCCCGGTGGCCCCGCAGCGCCTCCCGGTGGCCCCGCAGCGCCTCCCGGTGGCCGTCTTACCGTCCGA 574

Qy 401 GATCGTCCGCAACTGGGCGAGCTGTGCATGGGCGTGGCTAAGGCCGTGGTTGGTTTGGTTACG 460
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 573 GATCGTCCGCAACTGGGCGAGCTGTGCATGGGCGTGGCTAAGGCCGTGGTTGGTTTGGTTACG 514

Qy 461 ATTGGCCAGCGGGACTTAAGTGTGTCTCTGAAGAGCATGGACATTAGTGCTGGAGGGTCC 520
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 513 ATTGGCCAGCGGGACTTAAGTGTGTCTCTGAAGAGCATGGACATTAGTGCTGGAGGGTCC 454

Qy 521 TGGAAGAGTGATCCCCGCCCACCATCAAATGGGCGTTAGGTCTAGGAAGCGGGTGTGGG 580
       ||||||||||||||||||||||||||||||||||||||||||||||
Db 453 TGGAAGAGTGATCCCCGCCCACCATCAAATGGGCGTTAGGTTAGGAAGCGGGTGTGGG 394

Qy 581 TGGGGCCTTAGGGCGCAGAGGGCCAGACATACCCGAAGTGGTTGATTGTATACCGCAAGGG 640
       |||||||||||||||||||||||||||||||||||||||||||||
Db 393 TGGGGCCTTAGGGCGCAGAGGGCCAGACACACACCCGAAGTGGTTGATTGTATACCGCAAGGG 334

Qy 641 GCTGGATCGAACCCCCAAAGACACTGGAAGGCTGTGTGTGCTGAGGAGGGCCGGCA-AT 699
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 333 GCTGGATCGAACCCCCAAAGACACTGGAAGGCTGTGTGTGCTGAGGAGGGCCCGGCAGAT 274

Qy 700 CCAGTGTGTCGTGGGCCTTACAGGAAAGAGCTCCACCTTC-TTGGAGTGTGCAGATGGCGA 758
       |||||||||||||||||||||||||||||||||||||   ||||||||||||
Db 273 CCAGTGTGTCGTGGGCCTTTACAGGAAAGAGCTCCACCTTCTTTGGAGTGTGCAGATGTGCA 214

Qy 759 TCTAGGTGTGTCCACCCGATGGGAGCTGCGGGCCGGGCCGGCAGATGCTGCCCCAGTACAAAGC 818
       |||||||||||||||||||||||||||||||||||||||||||||||||
Db 213 TCTAGGTGTGTCCACCCGATGGGAGCTGCGGGCCGGGCCGGCAGATGCTGCCCCAGTACAAAGC 154

Qy 819 TGATTGGACCTGGGCCCTCTGGGCCTTCCCTGATTCTCTGCTTGCATCTCCAGCAAGTC 878
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 153 TGATTGGACCTGGGCCCTCTGGGCCTTCCCTGATTCTCTGCTTGCATCTCCAGCAAAGTC 94

Qy 879 CTGTCCCGTTGGCTGCCTTCATCCACTCTCTCACTTCTCTGCCTTCAGATAAATTGCA 938
       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 93  CTGTCCCGTTGGCTGCCTTCATCCACTCTCTCACTTCTCTGCCTTCAGAGTAAAATTGCA 34

Qy 939 AGATCTGTGTGCAAAAAAAAAAAAAAAAAA 971
       ||||||||||||||
Db 32  CAAGATCTGTGTGTGAAAAAAAAAAAAAAAA 1
```

```
RESULT 8
CA313142/c
LOCUS       CA313142           633 bp  mRNA   linear   EST   04-NOV-2002
DEFINITION  UI-CF-FN0-aex-n-23-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone
            UI-CF-FN0-aex-n-23-0-UI 3', mRNA sequence.
ACCESSION   CA313142
VERSION     CA313142.1  GI:24531240
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 633)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
```

Db   33 AGATCTGTGGTGCTTAAAAAAAAAAAAAAAAA 1

RESULT 9
CD742581/c
LOCUS   CD742581   623 bp   mRNA   linear   EST 26-JUN-2003
DEFINITION   UI-H-FT2-bjl-i-10-0-UI.s1 NCI_CGAP_FT2 Homo sapiens cDNA clone
              UI-H-FT2-bjl-i-10-0-UI 3', mRNA sequence.
ACCESSION   CD742581
VERSION   CD742581.1 GI:33293431
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 623)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Dr. Gary W. Hunninghake, U of I
              cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
              cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Distribution information can be found at
              http://genome.uiowa.edu/distribution/cgap.html
              Seq primer: M13 FORWARD
              POLYA=Yes.
FEATURES   Location/Qualifiers
     source   1..623
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-H-FT2-bjl-i-10-0-UI"
              /tissue_type="Aveolar Macrophage"
              /dev_stage="Adult"
              /lab_host="DH10B (Life Technologies)"
              /clone_lib="NCI_CGAP_FT2"
              /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
              modified polylinker; Site 1: EcoR I; Site 2: Not I;
              NCI_CGAP_FT2 is a subtracted cDNA library constructed from
              a pool of 81 RNA samples from Alveolar Macrophages
              challenged with different treatments. The library was
              subtracted according to Bonaldo, Lennon and Soares, Genome
              Research, 6:791-806, 1996. The tissue was provided by Dr.
              Gary W. Hunninghake of the University of Iowa.
              TAG_TISSUE=Human Lung Aveolar Macrophage
              TAG_LIB=UI-H-FT2
              TAG_SEQ=GCCATGCCG"
ORIGIN

Query Match   59.4%; Score 587.8; DB 14; Length 623;
Best Local Similarity 98.6%; Pred. No. 2.6e-79;
Matches   614; Conservative   0; Mismatches   7; Indels   2; Gaps   2;

Qy   343 CTCCCGTGGCCCCGCAGCGCCTCCGGTGCCCTGGCCTCTACCGTCCGAGA 402
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   623 CTCCCGTGGCCCCGCAGCGCCTCCGGTGCCCTGGCCTCTACCGTCCGAGA 564

Qy   403 TCGTCCCAACTGGGGGAGCTGTGCATGGGCGTGGCTAAGGCCGTGGTTTGGTTACGAT 462
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   563 TCGTCCCAACTGGGGGAGCTGTGCATGGGCGTGGCTAAGGCCGTGGTTTGGTTACGAT 504

Qy   463 TGGCCAGCGGGACTTAAGTGTTGTCTCTGAAGACATGGACATTAGTCTGGAGGGTCCTG 522
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   503 TGGCCAGCGGGACTTAAGTGTTGTCTCTGAAGACATGGACATTAGTCTGGAGGGTCCTG 444

Qy   523 GAAGAGTGATCCCCGCCCACCATCAAATGCGCGCTTAGGTCTAGGAAGCGGGGTGTGGGTG 582
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   443 GAAGAGTGATCCCCGCCCACCATCAAATGCGCGCTTAGGTCTAGGAAGCGGGGTGTGGGTG 384

Qy   583 GGGCCTTAGGGCGGGAGGGCCGCAGCATACCCCGAGTGGTTGGATTGTGTATACCGCAAGGGGC 642
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   383 GGGCCTTAGGGCGGGAGGGCCGCAGCACACCCCGAGTGGTTGGATTGTGTATACCGCAAGGGGC 324

Qy   643 TGGATCGAACCCCCAAAGACACTGGAAGGCTGGTGTGGCTGAGGAGGGCCCGGCCA-ATCC 701
         |||||||||||||||||||||||||||||||||||||||||| |||||||
Db   323 TGGATCGAACCCCCAAAGACACTGGAAGGCTGGTGTGGCTGAGGAGGGCCCGGCCAGATCC 264

Qy   702 AGTGTGTGTGGCTTACAGGAAAGAGTCCACCTCT-TGGAGTGTGCAGATGCGATC 760
         ||||||||||||||||||||||||||||||||||| ||||||||||||
Db   263 AGTGTGTGTGGCTTACAGGAAAGAGTCCACCTTCTCTGGAGTGTGCAGATGCGATC 204

Qy   761 TAGGTGTGTCCACCGATGGGAGCTGCGGCGCCGGCAGATGCGTGCCCCAGTACAAAGCTG 820
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203 TAGGTGTGTCCACCGATGGGAGCTGCGGCGCCGGCAGATGCGTGCCCCAGTACAAAGCTG 144

Qy   821 ATTTGGACTGGGGCCTTGGACTTCCCTGATTCTGCTTCTGCTTGCATCTCCAGCAAAGTCCT 880
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   143 ATTTGGACTGGGGCCTTCCTGGACTTCCCTGATTCTGCTTGCATCTCCAGCAAAGTCCT 84

Qy   881 GTCCCGTTGGCTGCTTCATCCACTCTGCCTTCTCTGCCTTCAGAGTAAAATTGCAAG 940
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   83 GTCCCGTTGGCTGCCTTCATCCACTCTCACTTCTCTGCCTTCAGAGTAAAATTGCAAG 24

Qy   941 ATCTGTGGTGCAAAAAAAAAAAA 963
         |||||||||
Db   23 ATCTGAAAAAAAAAAAAAAAAAA 1

RESULT 10
BQ447041/c
LOCUS   BQ447041   628 bp   mRNA   linear   EST 29-MAY-2002
DEFINITION   UI-H-EU1-bac-p-06-0-UI.s1 NCI_CGAP_Ctl Homo sapiens cDNA clone
              UI-H-EU1-bac-p-06-0-UI 3', mRNA sequence.
ACCESSION   BQ447041
VERSION   BQ447041.1 GI:21250153
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 628)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Dr. Jose Mercuende
              cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
              cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Clone distribution information can be obtained
              from Dr. M. Bento Soares, bento-soares@uiowa.edu
              Seq primer: M13 FORWARD
              POLYA=Yes.
FEATURES   Location/Qualifiers
     source   1..628
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-H-EU1-bac-p-06-0-UI"
              /tissue_type="Osteoarthritic Cartilage"
              /dev_stage="Adult"
              /lab_host="DH10B (Life Technologies)"
              /clone_lib="NCI_CGAP_Ctl"
              /note="Organ: Knee; Vector: pT7T3-Pac (Pharmacia) with a
              modified polylinker; Site 1: EcoR I; Site 2: Not I;
              NCI_CGAP_Ctl is a normalized cDNA library containing the
              following tissue(s): Osteoarthritic Cartilage The library
              was constructed according to Bonaldo, Lennon and Soares,
              Genome Research, 6:791-806, 1996. First strand cDNA
              synthesis was primed with an oligo-dT primer containing a

```
              Not I site. Double stranded cDNA was ligated to an EcoR I
              adaptor, digested with Not I, and cloned directionally
              into pT7T3-Pac vector. The oligonucleotide used to prime
              the synthesis of first-strand cDNA contains a library tag
              sequence that is located between the Not I site and the
              (dT)18 tail. The sequence tag for this library is
              TGATCACGCT.
              TAG_LIB=UI-H-EU1
              TAG_SEQ=TGATCACGCT"
ORIGIN

Query Match      58.2%;   Score 576.2;   DB 13;   Length 628;
Best Local Similarity  98.9%;   Pred. No. 1.5e-77;
Matches   622;   Conservative  0;   Mismatches  3;   Indels  4;   Gaps  4;

Qy   345  CCCGGTGGCCCCGCAGCGCCTCCCG-TGGCCTGGAGTGCAGAGTGCAGGTCTTACCGTCCGAGAT  403
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   628  CCCGGTGGCCCCGCAGCGCCTCCCGGTGGCCTGGAGTGCAGAGTGCAGGTCTTACCGTCCGAGAT  569

Qy   404  CGTCCGCAACTGGGCGGAGTGTGATGGGCGGCCTGGCTAAGGCCGTGGTTTGGTTACGATT  463
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568  CGTCCGCAACTGGGCGGAGCTGTGCATGGGCGGCCTGGCCGTGGTTTGGTTACGATT  509

Qy   464  GGCCAGCGGGACTTAAGTGTGTCTCTGAAGAGCATGGACATTAGTCTGGAGGTCCTG  523
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508  GGCCAGCGGGACTTAAGTGTGTCTCTGAAGAGCATGGACATTAGTCTGGAGGTCCTG  449

Qy   524  AAGAGTGATCCCGCCCACCATCAAATGGGCTTAGGTCTAGGAAGCGGGTGTGGGTGG  583
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   448  AAGAGTGATCCCGCCCACCATCAAATGGGCGCGCTTAGGTCTAGGTCTAGGTGTGGGTGG  389

Qy   584  GGCCTTAGGGCGGAGGCGCAGACATACCCCGAAGTGGTTGGATTGTATACCGGCAAGGGGCT  643
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388  GGCCTTAGGGCGGAGGCGCAGACACACCCCGAAGTGGTTGGATTGTATACCGGCAAGGGGCT  329

Qy   644  GGATCGAACCCCCAAAGACACTGGAAGGCTGTGTGGCTGAGGAGGGGCCGGCA-ATCCA  702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   328  GGATCGAACCCCCAAAGACACTGGAAGGCTGTGTGGCTGAGGAGGGGCCGGCCCGGCAGATCCA  269

Qy   703  GTGTGTCGTGGGCTTTACAGGAAAGAGCTTCCACCTTCT-TGGAGTGTGCAGATCGCGATCT  761
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   268  GTGTGTCGTGGGC-TTACAGGAAAGAGCTTCCACCTTCTCTCTGGAGTGTGCAGATCGCGATCT  210

Qy   762  AGGTGTGTCCACCCGATGGGGAGCTGCGGGGCCGGCGAGATGCTGCCCCAGTACAAAGCTGA  821
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   209  AGGTGTGTCCACCCGATGGGGAGCTGCGGGGCCGGCGAGATGCTGCCCCAGTACAAAGCTGA  150

Qy   822  TTTGGACCTGGGGCCTCTGGACTTCCCTGATTCTCTGCTTCCTGATTCTCCAGCAAGTCCTG  881
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   149  TTTGGACCTGGGGCCTCTGGACTTCCCTGATTCTCTGCTTCCTGATTCTCCAGCAAGTCCTG  90

Qy   882  TCCCGTTGGCTGCCTTCATCCCACTTCTCTGCCTTCAGAGTAAAATTGCAAGA  941
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    89  TCCCGTTGGCTGCCTTCATCCCACTTCTCTGCCTTCAGAGTAAAATTGCAAGA  30

Qy   942  TCTGTGTGGTGCAAAAAAAAAAAAAAAAAAAA  970
          ||||||||||||||||||||||||||||||||
Db    29  TCTGTGTGGTGCTAAAAAAAAAAAAAAAAAAA  1

RESULT 11
BF509528/c
LOCUS       BF509528     480 bp    mRNA    linear    EST 06-DEC-2000
DEFINITION  UI-H-B14-aoz-f-10-0-UI.s1 NCI_CGAP_Sub8 Homo sapiens cDNA clone
            IMAGE:3086731 3', mRNA sequence.
ACCESSION   BF509528
VERSION     BF509528.1  GI:11592826
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1 (bases 1 to 480)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            NCI-CGAP clone distribution information can be found through the
            I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..480
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3086731"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Sub8"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: Not I; Site 2: Eco RI; NCI_CGAP_Sub8
                     is a subtracted library derived from NCI_CGAP_Sub5. The
                     NCI_CGAP_Sub8 library had 2.5 million recombinants. A
                     single-stranded DNA preparation of NCI_CGAP_Sub5 was used
                     as a tracer in a subtractive hybridization with a driver
                     comprising: a pool of clones from NCI_CGAP_Sub5 (IMAGE
                     clone Ids 273283-2733415, 3068040-3069191; 25% of the
                     driver population), a pool of clones from NCI_CGAP_Sub4
                     (IMAGE clone Ids 272592-2729326; 25% of the driver
                     population), NCI_CGAP_Sub6 (pool AIF-AJU, IMAGE Ids
                     2728969-2733190; 25% of the driver population), and
                     NCI_CGAP_Sub7 (IMAGE Ids 3069192-3072238,
                     3081864-3084550; 25% of the driver population).
                     Subtraction was performed as previously described
                     [Bonaldo, Lennon & Soares (1996): Normalization and
                     Subtraction: Two Approaches To Facilitate Gene Discovery.
                     Genome Research 6, 791-806.
                     TAG_TISSUE=lung
                     TAG_LIB=NCI_CGAP_Lu19
                     TAG_SEQ=GACAGC"
ORIGIN

Query Match      45.4%;   Score 449.6;   DB 10;   Length 480;
Best Local Similarity  98.8%;   Pred. No. 2e-58;
Matches   474;   Conservative  0;   Mismatches  4;   Indels  2;   Gaps  2;

Qy   493  AGAGCATGACATTAGTCTGGAGGTCCTGGAGAGTGATCCCGGCCCCACCATCAAATG  552
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   480  AGAGCATGACATTAGTCTGGAGGTCCTGGAGGAGTGATCCCGCCCCACCATCAAATG  421

Qy   553  GGCTTAGGTCTAGGAAGCGGGTGTGGGTGGGGGGGCCTTAGGCGGAGGCGGCCAGACATACCCC  612
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   420  GGCTTAGGTCTAGGAAGCGGGTGTGGGTGGGGGGGCCTTAGGCGGAGGCGGCCAGACACACCCC  361

Qy   613  GAAGTGGTTGGATTGTATACCGGCAAGGGGCTGGATCGAACCCCCAAAGACACTGGAAGG  672
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   360  GAAGTGGTTGGATTGTATACCGGCAAGGGGCTGGATCGAACCCCCAAAGACACTGGAAGG  301

Qy   673  CTGTGTGGCTGAGGAGGGGCCGGCA-ATCAGTGTGTCGTGGGCTTTACAGGAAAGAGCT  731
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   300  CTGTGTGGCTGAGGAGGGGCCGGCAGATCCAGTGTGTCGTGGGCTTTACAGGAAAGAGCT  241

Qy   732  CCACCTTCT-TGGAGTGTGCAGATGCGATCTAGGTGTGTCCACCCGATGGGGAGCATACCCC  790
          ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   240  CCACCTTCTCTGGAGTGTGCAGATGCGATCTAGGTGTGTCCACCCGATGGGGAGCACACCCC  181

Qy   791  CCGGGCAGATGCTGCCCCAGTACAAAGCTGATTTGGACCTGGGGCCTCTGGACTTCCCTG  850
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Best Local Similarity　99.3%;　Pred. No. 3.3e-55;
Matches 451;　Conservative　0;　Mismatches　1;　Indels　2;　Gaps　2;

```
Qy   500 GGACATTAGTCTGGAGGTCCTGGAAGAGTGATCCCGCCCCATCAAATGGCGCTTA 559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     9 GGACATTAGTCTGGAGGTCCTGGAAGAGTGATCCCGCCCCATCAAATGGCGCTTA 68

Qy   560 GGTCTAGGAAGCGGGTGTGGGTGGGGCCTTAGGGCGGAGGCGCAGACATACCCGGAGTGG 619
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    69 GGTCTAGGAAGCGGGTGTGGGTGGGGCCTTAGGGCGGAGGCGCAGACACCCCGGAGTGG 128

Qy   620 TTGGATTGTATACCGGCAAGGGGCTGGATCGAACCCCCAAGACACTGGAGGGCTGTGTG 679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   129 TTGGATTGTATACCGGCAAGGGGCTGGATCGAACCCCCAAGACACTGGAGGGCTGTGTG 188

Qy   680 GCTGAGGAGGGCCCGGCA-ATCCAGTGTGTCGTGGGCTTTACAGGAAAGAGCTCCACCTT 738
             ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   189 GCTGAGGAGGGCCCGGCAGATCCAGTGTGTCGTGGGCTTTACAGGAAAGAGCTCCACCTT 248

Qy   739 CT-TGGAGTGTGCAGATGCGATCTAGGTGTGTCCACCCGATGGGAGCTGCGGGCCGGGCA 797
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   249 CTCTGGAGTGTGCAGATGCGATCTAGGTGTGTCCACCCGATGGGAGCTGCGGGCCGGGCA 308

Qy   798 GATGCTGCCCCAGTACAAAGCTGATTTGGACCTGGGCCTCTGGACTTCCCTGATTCTCT 857
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   309 GATGCTGCCCCAGTACAAAGCTGATTTGGACCTGGGCCTCTGGACTTCCCTGATTCTCT 368

Qy   858 GCTTGCATCTCCAGCAAAGTCCTGCCCGTTGGCTGCCTTCATCCACTCTCTCACTTCTC 917
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   369 GCTTGCATCTCCAGCAAAGTCCTGCCCGTTGGCTGCCTTCATCCACTCTCTCACTTCTC 428

Qy   918 TGCCTTCAGAGTAAAATTGCAAGATCTGTGGTGC 951
             ||||||||||||||||||||||||||||||||||
Db   429 TGCCTTCAGAGTAAAATTGCAAGATCTGTGGTGC 462
```

RESULT 13
BE677246/c
LOCUS　　　　BE677246　　　　　　439 bp　　　mRNA　　　linear　　EST 08-SEP-2000
DEFINITION　7d82e01.x1 Lupski_dorsal_root_ganglion Homo sapiens cDNA clone
　　　　　　 IMAGE:3279480 3', mRNA sequence.
ACCESSION　 BE677246
VERSION　　 BE677246.1　GI:10037787
KEYWORDS　　EST.
SOURCE　　　Homo sapiens (human)
ORGANISM　　Homo sapiens
　　　　　　 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　　 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE　 1　(bases 1 to 439)
AUTHORS　　 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE　　　 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
　　　　　　 Tumor Gene Index
JOURNAL　　 Unpublished (1997)
COMMENT　　 Contact: Robert Strausberg, Ph.D.
　　　　　　 Email: cgapbs-r@mail.nih.gov
　　　　　　 Tissue Procurement: Dr. James Lupski
　　　　　　 cDNA Library Preparation: Lupski Laboratory
　　　　　　 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
　　　　　　 DNA Sequencing by: Washington University Genome Sequencing Center
　　　　　　 Clone distribution: NCI-CGAP clone distribution information can be
　　　　　　 found through the I.M.A.G.E. Consortium/LLNL at:
　　　　　　 info@image.llnl.gov
　　　　　　 Seq primer: -40UP from Gibco
　　　　　　 High quality sequence stop: 413.
FEATURES　　　　　　　　　　　　 Location/Qualifiers
　　　source　　　　　　　 1..439
　　　　　　　　　　　　　 /organism="Homo sapiens"
　　　　　　　　　　　　　 /mol_type="mRNA"
　　　　　　　　　　　　　 /db_xref="taxon:9606"
　　　　　　　　　　　　　 /clone="IMAGE:3279480"
　　　　　　　　　　　　　 /sex="male"
　　　　　　　　　　　　　 /tissue_type="dorsal root ganglia"
　　　　　　　　　　　　　 /dev_stage="adult, 36 yr"

```
Db   180 CCGGGCCAGATGCTGCCCCAGTACAAAGCTGATTTGGACCTGGGCGCCTTCGGGGACTTCCCCTG 121
```

```
Qy   851 ATTCTCTGCTTGCATCTCCAGCAAAGTCCTGTCCCGTTGGCTGCCTTCATCCACTCTCTC 910
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   120 ATTCTCTGCTTGCATCTCCAGCAAAGTCCTGTCCCGTTGGCTGCCTTCATCCACTCTCTC 61

Qy   911 ACTTCTCTGCCTTCAGAGTAAAATTGCAAGATCTGTGGTGCAAAAAAAAAAAAAAAA 970
             |||||||||||||||||||||||||||||||||||| |||||||||||||||
Db    60 ACTTCTCTGCCTTCAGAGTAAAATTGCAAGATCTGTGGTGCTTAAAAAAAAAAAAAAA 1
```

RESULT 12
BM687366
LOCUS　　　　BM687366　　　　　462 bp　　mRNA　　　linear　　EST 28-FEB-2002
DEFINITION　UI-E-CQ1-ada-e-09-0-UI.r1 UI-B-CQ1 Homo sapiens cDNA clone
　　　　　　 UI-E-CQ1-ada-e-09-0-UI 5', mRNA sequence.
ACCESSION　 BM687366
VERSION　　 BM687366.1　GI:19000624
KEYWORDS　　EST.
SOURCE　　　Homo sapiens (human)
ORGANISM　　Homo sapiens
　　　　　　 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　　 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE　 1　(bases 1 to 462)
AUTHORS　　 Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE　　　 Normalization and subtraction: two approaches to facilitate gene
　　　　　　 discovery
JOURNAL　　 Genome Res. 6 (9), 791-806 (1996)
MEDLINE　　 97044477
PUBMED　　　8889548
COMMENT　　 Contact: Soares, MB
　　　　　　 Coordinated Laboratory for Computational Genomics
　　　　　　 University of Iowa
　　　　　　 375 Newton Road, 4156　MEBRF, Iowa City, IA 52242, USA
　　　　　　 Tel: 319 335 8250
　　　　　　 Fax: 319 335 9565
　　　　　　 Email: bento-soares@uiowa.edu
　　　　　　 Tissue Procurement: Dr. Gregg Hageman
　　　　　　 cDNA Library Preparation: Dr. M. Bento Soares, Univeristy of Iowa
　　　　　　 cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
　　　　　　 DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
　　　　　　 Clone Distribution: Researchers may obtain clones from Research
　　　　　　 Genetics (www.resgen.com).
　　　　　　 Seq primer: M13 Reverse.
FEATURES　　　　　　　　　　　　 Location/Qualifiers
　　　source　　　　　　　 1..462
　　　　　　　　　　　　　 /organism="Homo sapiens"
　　　　　　　　　　　　　 /mol_type="mRNA"
　　　　　　　　　　　　　 /db_xref="taxon:9606"
　　　　　　　　　　　　　 /clone="UI-E-CQ1-ada-e-09-0-UI"
　　　　　　　　　　　　　 /tissue_type="optic nerve"
　　　　　　　　　　　　　 /dev_stage="adult"
　　　　　　　　　　　　　 /lab_host="DH10B (life Technologies) (T1 phage resistant)"
　　　　　　　　　　　　　 /clone_lib="UI-E-CQ1"
　　　　　　　　　　　　　 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
　　　　　　　　　　　　　 modified polylinker; Site 1: EcoR I; Site 2: Not I;
　　　　　　　　　　　　　 UI-E-CQ1 is a normalized cDNA library containing the
　　　　　　　　　　　　　 following tissue(s): optic nerve. The library was
　　　　　　　　　　　　　 constructed according to Bonaldo, Lennon and Soares,
　　　　　　　　　　　　　 Genome Research, 6:791-806, 1996. First strand cDNA
　　　　　　　　　　　　　 synthesis was primed with an oligo-dT primer containing a
　　　　　　　　　　　　　 Not I site. Double stranded cDNA was ligated to an EcoR I
　　　　　　　　　　　　　 adaptor, digested with Not I, and cloned directionally
　　　　　　　　　　　　　 into pT7T3-Pac vector. The oligonucleotide used to prime
　　　　　　　　　　　　　 the synthesis of first-strand cDNA contains a library tag
　　　　　　　　　　　　　 sequence that is located between the Not I site and the
　　　　　　　　　　　　　 (dT)18 tail. This library was created for the program, Gene
　　　　　　　　　　　　　 CCATTAAGTG. Discovery in the Visual System, supported by National Eye
　　　　　　　　　　　　　 Institute (NEI)."

ORIGIN

Query Match　43.3%;　Score 428.4;　DB 12;　Length 462;

Seq primer: -40UP from Gibco
High quality sequence stop: 408.

FEATURES   Location/Qualifiers
source   1..411
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:1724708"
                /dev_stage="two placentae: one from 8 weeks and another
                from 9 weeks post conception"
                /lab_host="DH10B (ampicillin resistant)"
                /clone_lib="Soares placenta 8to9weeks 2NbHP8to9W"
                /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                strand cDNA was primed with a Not I - oligo(dT) primer [5',
                TGTTACCAATCTGAAGTGGAGCGGCCGCGATTTTTTTTTTTTTTTT 3'],
                double-stranded cDNA was size selected, ligated to Eco RI
                adapters (Pharmacia), digested with Not I and cloned into
                the Not I and Eco RI sites of a modified pT7T3 vector
                (Pharmacia). Library constructed by Bento Soares and
                M.Fatima Bonaldo."

ORIGIN

Query Match   38.7%;   Score 382.8; DB 9;   Length 411;
Best Local Similarity 98.8%;   Pred. No. 2.6e-48;
Matches   406; Conservative   0; Mismatches   3; Indels   2;   Gaps   2;

Qy   545 ATCAAATGCGCTTAGGTCTTAGGAAGCGGTGTGGGCCTTAGGGCGAGGCGCAGA 604
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   411 ATCAAATGCGCTTAGGTCTTAGGAAGCGGTGTGGGCCTTAGGGCGAGGCGCAGA 352

Qy   605 CATCCCCGAAGTGGTTGGATTGTATACCGCAAGGGGCTGGATCGAACCCCCAAAGACA 664
           ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   351 CACACNCCGAAGTGGTTGGATTGTATACCGCAAGGGGCTGGATCGAACCCCCAAAGACA 292

Qy   665 CTGGAAGGCTGTGTGGCTGAGGAGGGCCCGGCA-ATCCAGTGTGTCGTGGGCTTTACAGG 723
           ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   291 CTGGAAGGCTGTGTGGCTGAGGAGGGCCCGGCAGATCCAGTGTGTCGTGGGCTTTACAGG 232

Qy   724 AAAGAGCTCACCTTCT-TGGAGTGTGCAGATGCGATCTAGGTGTGTCACCCGATGGGGA 782
           |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   231 AAAGAGCTCACCTTCTTCTGGAGTGTGCAGATGCGATCTAGGTGTGTCACCCGATGGGGA 172

Qy   783 GCTGCGGGCCGGGCAGAGATGCTGCCCCAGTACAAAGCTGATTTGGACTCGGGGCTCTGGA 842
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   171 GCTGCGGGCCGGGCAGAGATGCTGCCCCAGTACAAAGCTGATTTGGACTCGGGGCTCTGGA 112

Qy   843 CTTCCCTGATTCTGCTTGCATCTCCAGCAAGTCCTGTCCCGTTGGCTGGCTTCATCC 902
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   111 CTTCCCTGATTCTGCTTGCATCTCCAGCAAGTCCTGTCCCGTTGGCTGGCTTCATCC 52

Qy   903 ACTCTCTCACTTCTCTGCCTTCTGCCTCCAGAGTAAAATTGCAAGATCTGTGGTGCAA 953
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   51   ACTCTCTCACTTCTCTGCCTTCTGCCTCCAGAGTAAAATTGCAAGATCTGTGGTGAAA 1

RESULT 15
AA687318/c
LOCUS   AA687318   395 bp   mRNA   linear   EST 24-DEC-1997
DEFINITION nv6le01.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1234296 3',
            mRNA sequence.
ACCESSION   AA687318
VERSION   AA687318.1 GI:2675509
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 395)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)

/lab_host="DH10B"
/clone_lib="Lupski dorsal_root_ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCCG-3' and
5'-GACTAGTTCGATACGGCAGCGGGGCCGCCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J.Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN

Query Match   41.4%;   Score 409.8; DB 10;   Length 439;
Best Local Similarity 99.1%;   Pred. No. 2.1e-52;
Matches   433; Conservative   0; Mismatches   0; Indels   2;   Gaps   2;

Qy   513 GAGGGTCCTGGAAGAGTGATCCCGGCCCACCATCAAATGCGCTTAGGTCTTAGGAAGCG 572
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   439 GAGGGTCCTGGAAGAGTGATCCCGGCCCACCATCAAATGCGCTTAGGTCTTAGGAAGCG 380

Qy   573 GGTGTGGGTGGGGGCCTTAGGGCGAGGCGCAGACATACCCCGAAGTGGTTGGATTGTATAC 632
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   379 GGTGTGGGTGGGGGCCTTAGGGCGAGGCGCAGACACACACCCCGAAGTGGTTGGATTGTATAC 320

Qy   633 CGCCAAGGGGCTGGATCGAACCCCCAAAGACACACTGGAAGGCTGTGTGGCTGAGGAGGGCC 692
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   319 CGCCAAGGGGCTGGATCGAACCCCCAAAGACACACTGGAAGGCTGTGTGGCTGAGGAGGGCC 260

Qy   693 CGGCA-ATCCAGTGTGTCGTGGGCTTCACCTTCT-TGGAGTGTGC 750
           ||||| ||||||||||||||||||||||||||||| |||||||||
Db   259 CGGCAGATCCAGTGTGTCGTGGGCTTTACACGGCTTTACACGGAGTGTGC 200

Qy   751 AGATGCGATCTAGGTGTGTCACCCGATGGGAGCTGCGGGCCGCTCCCCAG 810
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   199 AGATGCGATCTAGGTGTGTCACCCGATGGGAGCTGCGGGCCGCTGCCCCAG 140

Qy   811 TACAAAGCTGATTTGGACCTGGGCCTCGGCTGCTTCCTGATTCTCTGCTTGCATCTCCA 870
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   139 TACAAAGCTGATTTGGACCTGGGCCTCGGCTGCTTCCTGATTCTCTGCTTGCATCTCCA 80

Qy   871 GCAAAGTCCTGTCCCGTTGGCTGGCTTCATCCACTCTCTCACTTCTCTGCTTCAGAGTA 930
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79   GCAAAGTCCTGTCCCGTTGGCTGGCTTCATCCACTCTCTCACTTCTCTGCTTCAGAGTA 20

Qy   931 AAATTGCAAGATCTGTG 947
           |||||||||||||||||
Db   19   AAATTGCAAGATCTGTG 3

RESULT 14
AI184808/c
LOCUS   AI184808   411 bp   mRNA   linear   EST 28-OCT-1998
DEFINITION qd24g11.x1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA
            clone IMAGE:1724708 3', mRNA sequence.
ACCESSION   AI184808
VERSION   AI184808.1 GI:3735446
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 411)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 935   Std Error: 0.00

us-09-972-032-1_1.rst

Wed Jul 14 08:36:17 2004

```
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             unknown library type
             Insert Length: 966    Std Error: 0.00
             Seq primer: -40m13 fwd. ET from Amersham
             High quality sequence stop: 361.
FEATURES             Location/Qualifiers
     source          1..395
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1234296"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_GCB1"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGGCGGCCGCCTCATTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN

     Query Match         37.5%;  Score 371; DB 9;  Length 395;
     Best Local Similarity   99.5%;  Pred. No. 1.6e-46;
     Matches  393;  Conservative   0;  Mismatches   0;  Indels   2;  Gaps   2;

Qy    573 GGTGTGGGTGGGCCTTAGGGCGGAGGCGCAGACACATACCCCGAAGTGGTTGGATTGTATAC 632
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    395 GGTGTGGGTGGGCCTTAGGGCGGAGGCGCAGACACATACCCCGAAGTGGTTGGATTGTATAC 336

Qy    633 CGCAAGGGGCTGGATCGAACCCCCAAAGACACTGGAAGGCTGTGTGGCTGAGGAGGGGC 692
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    335 CGCAAGGGGCTGGATCGAACCCCCAAAGACACTGGAAGGCTGTGTGGCTGAGGAGGGGC 276

Qy    693 CGGCA-ATCCAGTGTGTGTGGGCTTTACAGGAAAGAGCTCCACCTTCT-TGGAGTGTGC 750
          ||||| ||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    275 CGGCAGATCCAGTGTGTGTGGGCTTTACAGGAAAGAGCTCCACCTTCTCTGGAGTGTGC 216

Qy    751 AGATGCGATCTAGGTGTGTCCACCCGATGGGAGCTGCGGGCCGGGCAGATGCTGCCCCAG 810
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    215 AGATGCGATCTAGGTGTGTCCACCCGATGGGAGCTGCGGGCCGGGCAGATGCTGCCCCAG 156

Qy    811 TACAAAGCTGATTTGGACCTGGGGCCTCCTGGACTTCCTGATTCTCTGCTTGCATCTCCA 870
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    155 TACAAAGCTGATTTGGACCTGGGGCCTCCTGGACTTCCCTGATTCTCTGCTTGCATCTCCA 96

Qy    871 GCAAAGTCCTGTCCCCGTTGGCCTGCCTTCATCCACTTCTCCACTTCTCCGCCTTCAGAGTA 930
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     95 GCAAAGTCCTGTCCCCGTTGGCCTGCCTTCATCCACTTCTCCACTTCTCCGCCTTCAGAGTA 36

Qy    931 AAATTGCAAGATCTGGTGGTGCAAAAAAAAAAAAAA 965
          |||||||||||||||||||||||||||||||||||
Db     35 AAATTGCAAGATCTGGTGGTGCAAAAAAAAAAAAAA 1
```

Search completed: July 10, 2004, 11:56:41
Job time : 3119 secs